Radon Series Comp. Appl. Math **8**, 371–410

Optimal investment and hedging under partial and inside information

Michael Monoyios

Abstract. This article concerns optimal investment and hedging for agents who must use trading strategies which are adapted to the filtration generated by asset prices, possibly augmented with some inside information related to the future evolution of an asset price. The price evolution and observations are taken to be continuous, so the partial (and, when applicable, inside) information scenario is characterised by asset price processes with an unknown drift parameter, which is to be filtered from price observations. With linear observation and signal process dynamics, this is done with a Kalman–Bucy filter. Using the dual approach to portfolio optimisation, we solve the Merton optimal investment problem when the agent does not know the drift parameter of the underlying stock. This is taken to be a random variable with a Gaussian prior distribution, which is updated via the Kalman filter. This results in a model with a stochastic drift process adapted to the observation filtration, and which can be treated as a full information problem, yielding an explicit solution. We also consider the same problem when the agent has noisy knowledge at time zero of the terminal value of the Brownian motion driving the stock. Using techniques of enlargement of filtration to accommodate the insider's additional knowledge, followed by filtering the asset price drift, we are again able to obtain an explicit solution. Finally we treat an incomplete market hedging problem. A claim on a non-traded asset is hedged using a correlated traded asset. We summarise the full information case, then treat the partial information scenario in which the hedger is uncertain of the true values of the asset price drifts. After filtering, the resulting problem with random drifts is solved in the case that each asset's prior distribution has the same variance, resulting in analytic approximations for the optimal hedging strategy.

Key words. Duality, filtering, incomplete information, optimal portfolios.

AMS classification. 49N30, 93E11, 93C41, 91B28

1 Introduction

This article examines some problems of optimal investment, and of optimal hedging of a contingent claim in an incomplete market, when the agent's information set is restricted to stock price observations, possibly augmented by some additional information related to the terminal value of a stock price.

In classical models of financial mathematics, one usually specifies a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, and then writes down some stochastic process $S = (S_t)_{0\leq t\leq T}$ for an asset price, such that S is adapted to the filtration $\mathbb{F}$. A typical example would be the Black–Scholes (henceforth, BS) model of

Work presented during the *Special Semester on Stochastics with Emphasis on Finance*, September 3 – December 5, 2008, organised by RICAM (Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences) Linz, Austria

a stock price, following the geometric Brownian motion

$$\mathrm{d}S_t = \sigma S_t(\lambda \mathrm{d}t + \mathrm{d}B_t), \tag{1.1}$$

where B is a $(P, \mathbb{F})$-Brownian motion and the volatility $\sigma > 0$ and the Sharpe ratio λ are assumed to be known constants. Of course, this is a strong assumption that an agent is assumed to be able to observe the Brownian motion process B, as well as the stock price process S. We refer to this as a *full information* scenario. In this case, an agent uses $\mathbb{F}$-adapted trading strategies in S, a process with known drift and diffusion coefficients.

We shall frequently relax the full information assumption in this article. We shall assume that the agent can only observe the stock price process, and not the Brownian motion B. The agent's trading strategies must therefore be adapted to the *observation filtration* $\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_t)_{0\le t\le T}$ generated by S. This is a *partial information* scenario. In recent years there has been a growing research activity in this area, as surveyed by Pham [28], for instance, who examines some different scenarios to the ones in this article.

With partial information, the parameter λ would be regarded as an unknown constant whose value needs to be determined from price data. In principle, one would also have to apply this philosophy to the volatility σ, but we shall make the approximation that price observations are continuous, so that σ can be computed from the quadratic variation $[S]_t$ of the stock price, since we have

$$[S]_t = \sigma^2 S_t^2 t, \quad 0 \le t \le T.$$

One way to model the uncertainty in our knowledge of the (supposed constant) parameter λ is to take a so-called Bayesian approach. This means we consider λ to be an $\mathcal{F}_0$-measurable random variable with a given initial distribution (the *prior* distribution). The prior distribution initialises the probability law of λ conditional on $\widehat{\mathcal{F}}_0$, and this is updated in the light of new price information, that is, as the observation filtration $\widehat{\mathbb{F}}$ evolves. (In the case that λ is some unknown process $(\lambda_t)_{0\le t\le T}$ (as opposed to an unknown constant), then we would consider it to be some $\mathbb{F}$-adapted process such that its starting value λ_0 has a given prior distribution conditional on $\widehat{\mathcal{F}}_0$.)

This is an example of a *filtering* problem. In the case of the BS model (1.1), where we model λ as an $\mathcal{F}_0$-measurable random variable, we are interested in computing the conditional expectation

$$\widehat{\lambda}_t := E\big[\lambda \mid \widehat{\mathcal{F}}_t\big], \quad 0 \le t \le T.$$

We shall see that the effect of filtering is that the model (1.1) may be replaced by a model specified on the filtered probability space $(\Omega, \widehat{\mathcal{F}}_T, \widehat{\mathbb{F}}, P)$ and written as

$$\mathrm{d}S_t = \sigma S_t(\widehat{\lambda}_t \mathrm{d}t + \mathrm{d}\widehat{B}_t),$$

where $\widehat{B}$ is a $(P, \widehat{\mathbb{F}})$-Brownian motion. This model may now be treated as a full information model, since both $\widehat{B}$ and $\widehat{\lambda}$ are $\widehat{\mathbb{F}}$-adapted processes. The price we have paid for restoring a full information scenario is that the constant parameter λ has been replaced by a random process $\widehat{\lambda}$. The procedure by which a partial information model is replaced with a tractable full information model under the observation filtration is typically only achievable in special circumstances, such as Gaussian prior distributions and certain linearity properties in the relation between the observable and unobservable processes.

The rest of the article is as follows. In Section 2, we briefly introduce the innovations process of filtering theory and state the filtering algorithm that we shall use, the celebrated Kalman–Bucy filter [11]. In Section 3 we use the dual approach to portfolio optimisation (see Karatzas [13] for example), to solve the Merton problem [19, 20] of optimal investment, when the drift parameter of the stock must be filtered from price observations. In Section 4 we solve the Merton problem when the agent is again uncertain of the stock's drift, but is assumed to have some additional information in the form of knowledge of the value of a random variable I, representing noisy information on the underlying Brownian motion at time T. Further examples of optimal investment problems with inside information and parameter uncertainty are given in Danilova, Monoyios and Ng [2]. Finally, in Section 5 we consider the hedging of a claim in an incomplete market setting under partial information. Specifically, we consider a basis risk model involving the optimal hedging of a claim on a non-tradeable asset Y using a traded stock S, correlated with Y, when the hedger is restricted to trading strategies in S that are adapted to the observation filtration generated by the asset prices. A number of papers, such as Henderson [8], Monoyios [21, 22] and Musiela and Zariphopoulou [26], have used exponential indifference valuation methods to hedge the claim in an optimal manner in a full information scenario. We outline these results before moving on to the partial information case, where we assume the hedger does not know with certainty the drifts of S and Y. Analytic approximations for prices and hedging strategies are given. Further work on this topic can be found in Monoyios [23, 25].

2 Innovations and linear filtering

Filtering problems concern estimating something (in a manner to be made precise shortly) about an unobserved stochastic process Ξ given observations of a related process Λ. The problem was solved for linear systems in continuous time by Kalman and Bucy [11]. Subsequent work sought generalisations to systems with nonlinear dynamics, see Zakai [33] for instance. Kailath [10] developed the so-called innovations approach to linear filtering, which formulated the problem in the context of martingale theory. This approach to nonlinear filtering was given a definitive treatment by Fujisaki, Kallianpur and Kunita [7]. Textbook treatments can be found in Kallianpur [12], Lipster and Shiryaev [16, 17], Rogers and Williams [32], Chapter VI.8, and Fleming and Rishel [6].

The setting is a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$. All processes are assumed to be $\mathbb{F}$-adapted. Note that $\mathbb{F}$ is not the observation filtration. Let us call $\mathbb{F}$ the *background filtration*. We consider two processes, both taken to be one-dimensional (for simplicity):

- a *signal process* $\Xi = (\Xi_t)_{0 \le t \le T}$ which is not directly observable;
- an *observation process* $\Lambda = (\Lambda_t)_{0 \le t \le T}$, which is observable and somehow correlated with Ξ, so that by observing Λ we can say something about the distribution of Ξ.

Let $\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_t)_{0 \le t \le T}$ denote the observation filtration generated by Λ. That is,

$$\widehat{\mathcal{F}}_t := \sigma(\Lambda_s; 0 \le s \le t), \quad 0 \le t \le T.$$

The filtering problem is to compute the conditional expectation

$$\widehat{\Xi}_t := E\big[\Xi_t \mid \widehat{\mathcal{F}}_t\big], \quad 0 \le t \le T. \tag{2.1}$$

To proceed further, we specify some particular model for the observation and signal processes. We shall focus on the linear case where both Λ and Ξ are solutions to linear stochastic differential equations (SDEs).

2.1 Linear observations and linear signal

Let $B = (B_t)_{0 \le t \le T}$ be an $\mathbb{F}$-Brownian motion. We assume the observation process Λ is of the form

$$\Lambda_t = \int_0^t G(s)\Xi_s \mathrm{d}s + B_t, \quad 0 \le t \le T, \quad \text{(linear observation)} \tag{2.2}$$

with $G(\cdot)$ a deterministic function such that $E\int_0^T G^2(t)\Xi_t^2 < \infty$.

We take the signal process to be of the form

$$\Xi_t = \Xi_0 + \int_0^t A(s)\Xi_s \mathrm{d}s + \int_0^t C(s)\mathrm{d}W_s, \quad 0 \le t \le T, \quad \text{(linear signal)}$$

for deterministic functions $A(\cdot), C(\cdot)$, with W a $(P, \mathbb{F})$-Brownian motion independent of the $\mathcal{F}_0$-measurable random variable Ξ_0, and correlated with B in the observation model (2.2) according to

$$[W, B]_t = \rho t, \quad 0 \le t \le T, \quad \rho \in [-1, 1].$$

Suppose further that the signal process has a Gaussian initial distribution, $\Xi_0 \sim \mathrm{N}(\mu, \mathrm{v})$, independent of B and of W, where $\mathrm{N}(\mu, \mathrm{v})$ denotes the normal probability law with mean μ and variance v. The two-dimensional process (Ξ, Λ) is then Gaussian, so the conditional distribution of Ξ_t given the sigma-field $\widehat{\mathcal{F}}_t$ will also be normal (and so, in particular, is completely characterised by its mean and variance), with mean given by (2.1) and variance

$$V_t := \mathrm{var}\big[\Xi_t \mid \widehat{\mathcal{F}}_t\big] = E\big[(\Xi_t - \widehat{\Xi}_t)^2 \mid \widehat{\mathcal{F}}_t\big] = \widehat{\Xi_t^2} - \Big(\widehat{\Xi}_t\Big)^2, \quad 0 \le t \le T.$$

Notice that the initial values are

$$\widehat{\Xi}_0 = E\big[\Xi_0 \,|\, \widehat{\mathcal{F}}_0\big] = E\Xi_0 = \mu,$$

and

$$V_0 = E\big[(\Xi_0 - \widehat{\Xi}_0)^2 \,|\, \widehat{\mathcal{F}}_0\big] = E\big[(\Xi_0 - \mu)^2\big] = \mathrm{var}(\Xi_0) = \mathrm{v}.$$

The problem then boils down to finding an algorithm for computing the sufficient statistics $\widehat{\Xi}_t, V_t$ from their initial values $\widehat{\Xi}_0 = \mu$, $V_0 = \mathrm{v}$. For linear systems it turns out that the conditional variance V_t is a *deterministic* function of t. Thus, there is in fact only one sufficient statistic, the conditional mean $\widehat{\Xi}_t$, which turns out to satisfies a linear SDE. This is the celebrated *Kalman–Bucy filter*, given in Theorem 2.1 shortly.

2.2 Innovations process

Define the $\widehat{\mathbb{F}}$-adapted *innovations process* $N = (N_t)_{0 \le t \le T}$ by

$$N_t := \Lambda_t - \int_0^t G(s)\widehat{\Xi}_s \mathrm{d}s, \quad 0 \le t \le T.$$

We recall two crucial properties of the innovations process, which form the bedrock of filtering theory.

- The innovations process N is an $\widehat{\mathbb{F}}$-Brownian motion.
- Every local $\widehat{\mathbb{F}}$-martingale M admits a representation of the form

$$M_t = M_0 + \int_0^t \Phi_s \mathrm{d}N_s, \quad 0 \le t \le T,$$

  where Φ is $\widehat{\mathbb{F}}$-adapted and $\int_0^T \Phi_t^2 \mathrm{d}t < \infty$ a.s.

For a proof of the above results, and of the following celebrated result for filtering of linear systems, see Rogers and Williams [32] or Lipster and Shiryaev [16], for instance.

Theorem 2.1 (One-dimensional Kalman–Bucy filter). *On a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, with $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$, let $\Xi = (\Xi_t)_{0 \le t \le T}$ be an $\mathbb{F}$-adapted signal process satisfying*

$$\mathrm{d}\Xi_t = A(t)\Xi_t \mathrm{d}t + C(t)\mathrm{d}W_t,$$

and let $\Lambda = (\Lambda_t)_{0 \le t \le T}$ be an $\mathbb{F}$-adapted observation process satisfying

$$\mathrm{d}\Lambda_t = G(t)\Xi_t \mathrm{d}t + \mathrm{d}B_t, \quad \Lambda_0 = 0,$$

where W, B are $\mathbb{F}$-Brownian motions with correlation ρ, and the coefficients $A(\cdot)$, $C(\cdot)$, $G(\cdot)$ are deterministic functions satisfying

$$\int_0^T \big(|A(t)| + C^2(t) + G^2(t)\big)\,\mathrm{d}t < \infty.$$

Define the observation filtration $\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_t)_{0 \le t \le T}$ *by*

$$\widehat{\mathcal{F}}_t := \sigma(\Lambda_s; 0 \le s \le t).$$

Suppose Ξ_0 *is an* $\mathcal{F}_0$*-measurable random variable, and that the distribution of* Ξ_0 *is Gaussian with mean* μ *and variance* v*, independent of* W *and* B*. Then the conditional expectation* $\widehat{\Xi}_t := E\big[\Xi_t \mid \widehat{\mathcal{F}}_t\big]$*, for* $0 \le t \le T$*, satisfies*

$$\mathrm{d}\widehat{\Xi}_t = A(t)\widehat{\Xi}_t \mathrm{d}t + [G(t)V_t + \rho C(t)]\,\mathrm{d}N_t, \quad \widehat{\Xi}_0 = \mu,$$

where $N = (N_t)_{0 \le t \le T}$ *is the innovations process, an* $\widehat{\mathbb{F}}$*-Brownian motion satisfying*

$$\mathrm{d}N_t = \mathrm{d}\Lambda_t - G(t)\widehat{\Xi}_t \mathrm{d}t,$$

and $V_t = \mathrm{var}\big[\Xi_t \mid \widehat{\mathcal{F}}_t\big]$*, for* $0 \le t \le T$*, is the conditional variance, which is independent of* $\widehat{\mathcal{F}}_t$ *and satisfies the deterministic Riccati equation*

$$\frac{\mathrm{d}V_t}{\mathrm{d}t} = (1-\rho^2)C^2(t) + 2\big[A(t) - \rho C(t)G(t)\big]V_t - G^2(t)V_t^2, \quad V_0 = \mathrm{v}.$$

A multi-dimensional version of the Kalman–Bucy filter can be derived using similar techniques to the one-dimensional case. See Theorem V.9.2 in Fleming and Rishel [6], for instance.

3 Optimal investment problems with random drift

3.1 Portfolio optimisation via convex duality

We wish to apply the filtering results in the previous section to portfolio optimisation and optimal hedging problems when the agent does not know the drift parameters of the underlying assets. The filtering approach leads to portfolio problems in which the assets follow SDEs with random drift parameters. The dual approach to portfolio optimisation is now a classical technique, well suited to such problems. In this section we recall the main results of portfolio optimisation via convex duality. See Karatzas [13] for more details and further references.

Consider an agent with a continuous, differentiable, increasing, concave utility function $U : \mathbb{R}^+ \to \mathbb{R}$. Define the *convex conjugate* $\widetilde{U} : \mathbb{R}^+ \to \mathbb{R}$ of U by

$$\widetilde{U}(\eta) := \sup_{x \in \mathbb{R}^+} [U(x) - x\eta], \quad \eta > 0. \tag{3.1}$$

Then $\widetilde{U}$ is a decreasing, continuously differentiable, convex function given by

$$\widetilde{U}(\eta) = U(I(\eta)) - \eta I(\eta), \tag{3.2}$$

where I is the inverse of U'. Differentiating (3.2) gives

$$\widetilde{U}'(\eta) = -I(\eta). \tag{3.3}$$

We note that the defining duality relation (3.1) is equivalent to the bidual relation

$$U(x) = \inf_{\eta\in\mathbb{R}^+}[\widetilde{U}(\eta) + x\eta], \quad x > 0.$$

We are interested in solving an optimal portfolio problem for an agent in a complete market with a single stock whose price process is a continuous semimartingale. To be precise, on an a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$, suppose a stock price $S = (S_t)_{0\le t\le T}$ follows

$$\mathrm{d}S_t = \sigma_t S_t(\lambda_t \mathrm{d}t + \mathrm{d}B_t),$$

where $\sigma = (\sigma_t)_{0\le t\le T}$ and $\lambda = (\lambda_t)_{0\le t\le T}$ are $\mathbb{F}$-adapted processes, and $B = (B_t)_{0\le t\le T}$ is an $\mathbb{F}$-Brownian motion. For simplicity, we take the interest rate to be zero.

The wealth process $X = (X_t)_{0\le t\le T}$ associated with a self-financing portfolio involving S is given by

$$\mathrm{d}X_t = \sigma_t\theta_t X_t(\lambda_t\mathrm{d}t + \mathrm{d}B_t), \quad X_0 = x,$$

where the process $\theta = (\theta_t)_{0\le t\le T}$ represents the proportion of wealth placed in the stock, and constitutes the agent's trading strategy. Define the set $\mathcal{A}$ of admissible trading strategies as those satisfying $\int_0^T \sigma_t^2\theta_t^2\mathrm{d}t < \infty$ a.s. and whose wealth process satisfies $X_t \ge 0$ a.s. for all $t \in [0, T]$.

The unique martingale measure $Q \sim P$ on $\mathcal{F}_T$ is defined by

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = Z_T,$$

where $Z = (Z_t)_{0\le t\le T}$ is the exponential local martingale defined by

$$Z_t := \mathcal{E}(-\lambda \cdot B)_t, \quad 0 \le t \le T.$$

We assume that λ satisfies the Novikov condition

$$E\exp\left(\frac{1}{2}\int_0^T \lambda_t^2\mathrm{d}t\right) < \infty,$$

so that Z is indeed a martingale and Q is indeed a probability measure equivalent to P.

Under Q, the process B^Q defined by

$$B_t^Q := B_t + \int_0^t \lambda_s\mathrm{d}s, \quad 0 \le t \le T,$$

is a Brownian motion. The Q-dynamics of S, X are

$$\mathrm{d}S_t = \sigma_t S_t\mathrm{d}B_t^Q, \quad \mathrm{d}X_t = \sigma_t\theta_t X_t\mathrm{d}B_t^Q.$$

In particular, the solution of the SDE for X, given $X_0 = x$, is

$$X_t = x\mathcal{E}(\sigma\theta \cdot B^Q)_t, \quad 0 \le t \le T.$$

We assume that

$$E^Q \exp\left(\frac{1}{2}\int_0^T \sigma_t^2\theta_t^2 \mathrm{d}t\right) < \infty,$$

so that X is a Q-martingale, satisfying $E^Q X_T = x$, or

$$E\big[Z_T X_T\big] = x, \tag{3.4}$$

which we shall regard as a *constraint* on the terminal wealth X_T. This is the foundation of the dual approach to portfolio optimisation, namely to enforce the martingale constraint on the wealth process.

The basic portfolio problem (the *primal* problem) is, given $X_0 = x$, to maximise expected utility of wealth at time T:

$$u(x) := \sup_{\theta\in\mathcal{A}} EU(X_T), \tag{3.5}$$

subject to (3.4).

The *dual* value function is $\tilde{u} : \mathbb{R}^+ \to \mathbb{R}$ defined by

$$\tilde{u}(\eta) := E\widetilde{U}\left(\eta\frac{\mathrm{d}Q}{\mathrm{d}P}\right), \quad \eta > 0.$$

The well-known result on portfolio optimisation via duality for this model is as follows.

Theorem 3.1. *1. The primal and dual value functions $u(x)$ and $\tilde{u}(\eta)$ are conjugate:*

$$\tilde{u}(\eta) = \sup_{x\in\mathbb{R}^+}[u(x) - x\eta], \qquad u(x) = \inf_{\eta>0}[\tilde{u}(\eta) + x\eta],$$

so that $u'(x) = \eta$ (equivalently, $\tilde{u}'(\eta) = -x$);

2. The optimal terminal wealth in (3.5) is X_T^* satisfying

$$U'(X_T^*) = \eta\frac{\mathrm{d}Q}{\mathrm{d}P}, \quad equivalently, \quad X_T^* = I\left(\eta\frac{\mathrm{d}Q}{\mathrm{d}P}\right).$$

A proof of this result can be found in Karatzas [13]. The idea behind the proof is to consider the maximisation of the objective functional $EU(X_T)$ subject to the constraint $E\big[Z_T X_T\big] = x$, via the Lagrangian

$$L(X_T, \eta) := EU(X_T) + \eta\left(x - E\big[Z_T X_T\big]\right).$$

The first order condition for an optimum then yields that the optimal terminal wealth is characterised by

$$U'(X_T^*) = \eta Z_T \Leftrightarrow X_T^* = I\left(\eta Z_T\right). \tag{3.6}$$

The value of the multiplier η is needed to fully determine X_T^*. We substitute (3.6) into the constraint $E\big[Z_T X_T^*\big] = x$, so that η is given by

$$E\big[Z_T I(\eta Z_T)\big] = x,$$

or, using the definition of $\tilde{u}(\eta)$ and (3.3),

$$\tilde{u}'(\eta) = -x.$$

This is precisely the relation we expect to hold when u and $\tilde{u}$ are conjugate.

3.1.1 Duality for incomplete markets

Similar duality theorems have been developed for incomplete market situations, and also when the agent has a random terminal endowment, possibly in the form of a contingent claim. For the incomplete market case, see the seminal paper by Karatzas et al. [14] for markets with continuous price processes, and Kramkov and Schachermayer [15] for the case with general semimartingale price processes. For problems involving a terminal random endowment in the form of an $\mathcal{F}_T$-measurable random variable, contributions have been made by (among others) Hugonnier and Kramkov [9], Owen [27] and by Delbaen et al. [5] for an agent with an exponential utility function. We shall use the results of [5] in Section 5, when we examine the exponential hedging of a contingent claim in a basis risk model.

For an incomplete market, in which the set $\mathcal{M}$ of martingale measures is no longer a singleton, the significant change is that the dual value function is then defined by

$$\tilde{u}(\eta) := \inf_{Q \in \mathcal{M}} E\widetilde{U}\left(\eta \frac{\mathrm{d}Q}{\mathrm{d}P}\right). \tag{3.7}$$

The form of the duality theorem for an incomplete market is similar to Theorem 3.1, but with the unique martingale measure Q of the complete market replaced by the optimal *dual minimiser* Q^* that achieves the infimum in (3.7). See [13], for instance, for details in an Itô process setting.

3.2 Optimal investment with Gaussian drift process

We wish to apply filtering theory and the martingale approach to portfolio optimisation to the classical optimal portfolio problem of Merton [19, 20], in the case that the agent does not know the drift parameter of the stock. As we shall see, this will involve a portfolio problem in which the market price of risk of the stock is a Gaussian process. Hence we first describe the solution to such a problem.

Suppose a stock price $S = (S_t)_{0 \le t \le T}$ follows the process

$$\mathrm{d}S_t = \sigma S_t(\lambda_t \mathrm{d}t + \mathrm{d}B_t),$$

on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}, P)$, with B an $\mathbb{F}$-Brownian motion and λ an $\mathbb{F}$-adapted process following

$$\lambda_t = \lambda_0 + \int_0^t \mathrm{w}_s dB_s, \quad \mathrm{w}_t = \frac{\mathrm{w}_0}{1 + \mathrm{w}_0 t}, \quad 0 \le t \le T, \tag{3.8}$$

for constants λ_0, w_0.

The self-financing wealth process X from trading S is given by

$$\mathrm{d}X_t = \sigma\theta_t X_t(\lambda_t \mathrm{d}t + \mathrm{d}B_t), \quad X_0 = x, \tag{3.9}$$

where the trading strategy $\theta = (\theta_t)_{0 \le t \le T}$ is the proportion of wealth invested in stock. We define the set $\mathcal{A}$ of admissible strategies as those satisfying $\int_0^T \theta_t^2 \mathrm{d}t < \infty$ almost surely, such that $X_t \ge 0$ almost surely for all $t \in [0, T]$.

The value function is

$$u(x) := \sup_{\theta\in\mathcal{A}} E\big[U(X_T) \mid \mathcal{F}_0\big] \tag{3.10}$$

where $U(x)$ is the power utility function given by

$$U(x) = \frac{x^\gamma}{\gamma}, \quad 0 < \gamma < 1. \tag{3.11}$$

Theorem 3.2. *Assume that*

$$-1 < \mathrm{w}_0 T < \frac{1-\gamma}{\gamma}.$$

Then the value function (3.10) is given by

$$u(x) = \frac{x^\gamma}{\gamma} C^{1-\gamma}, \tag{3.12}$$

where C is given by

$$C = \left(\frac{(1+\mathrm{w}_0 T)^q}{1+q\mathrm{w}_0 T}\right)^{1/2} \exp\left(-\frac{1}{2}\frac{q(1-q)\lambda_0^2 T}{1+q\mathrm{w}_0 T}\right), \quad q = -\frac{\gamma}{1-\gamma}. \tag{3.13}$$

The optimal trading strategy θ^* achieving the supremum in (3.10) is given by

$$\theta_t^* = \frac{\lambda_t}{\sigma(1-\gamma)}\left(\frac{1}{1+q\mathrm{w}_t(T-t)}\right), \quad 0 \le t \le T. \tag{3.14}$$

Proof. Let Q denote the unique martingale measure for this market. The change of measure martingale $Z := (Z_t)_{0\le t\le T}$ is given by

$$Z_t := \left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \mathcal{E}(-\lambda\cdot B)_t, \quad 0 \le t \le T,$$

and satisfies the SDE

$$\mathrm{d}Z_t = -\lambda_t Z_t \mathrm{d}B_t, \quad Z_0 = 1. \tag{3.15}$$

Notice that

$$\lim_{\mathrm{w}_0\to 0} Z_t = \mathcal{E}(-\lambda_0 B)_t = \exp\left(-\lambda_0 B_t - \frac{1}{2}\lambda_0^2 t\right). \tag{3.16}$$

We may write $Z_t = f(t,\lambda_t)$ where $f : [0,T]\times\mathbb{R} \to \mathbb{R}^+$ is a smooth function, and apply Itô's formula along with the SDE (3.8) for λ to give

$$\mathrm{d}Z_t = \Big[f_t(t,\lambda_t) + \frac{1}{2}\mathrm{w}_t^2 f_{xx}(t,\lambda_t)\Big]\mathrm{d}t + \mathrm{w}_t f_x(t,\lambda_t)\mathrm{d}B_t, \tag{3.17}$$

with subscripts of f denoting partial derivatives. Equating (3.15) and (3.17) yields the partial differential equations for f:

$$\begin{aligned} \mathrm{w}_t f_x(t,x) &= -x f(t,x), \\ f_t(t,x) + \frac{1}{2}\mathrm{w}_t^2 f_{xx}(t,x) &= 0, \end{aligned}$$

with $f(0,\cdot) = Z_0 = 1$. The solution to these equations gives Z_t in the form

$$Z_t = \left(\frac{\mathrm{w}_0}{\mathrm{w}_t}\right)^{1/2} \exp\left[-\frac{1}{2}\left(\frac{\lambda_t^2}{\mathrm{w}_t} - \frac{\lambda_0^2}{\mathrm{w}_0}\right)\right], \quad 0 \le t \le T. \tag{3.18}$$

Note that this function is actually well-defined even for $\mathrm{w}_0 \to 0$. It is not hard to check that (3.18) reduces to (3.16) in the limit $\mathrm{w}_0 \to 0$.

For power utility, the convex conjugate $\widetilde{U}$ of the utility function is given by

$$\widetilde{U}(\eta) = -\frac{\eta^q}{q}, \quad q = -\frac{\gamma}{1-\gamma}, \quad \eta > 0. \tag{3.19}$$

The dual value function is defined by

$$\tilde{u}(\eta) := E\big[\widetilde{U}(\eta Z_T) \,|\, \mathcal{F}_0\big], \quad \eta > 0.$$

Using (3.19) we obtain

$$\tilde{u}(\eta) = -\frac{\eta^q}{q} C,$$

where

$$C := E\big[Z_T^q \,|\, \mathcal{F}_0\big]. \tag{3.20}$$

From Theorem 3.1, the primal and dual value functions are conjugate, which yields that the primal value function is indeed given by (3.12), with C defined by (3.20). It therefore remains to show that C is indeed equal to the expression in (3.13) and that the optimal strategy is given by (3.14).

Once again using Theorem 3.1, the optimal terminal wealth X_T^*, attained by adopting the strategy that achieves the supremum in (3.10), is given by

$$X_T^* = -\widetilde{U}'(u'(x) Z_T).$$

Hence, using the form (3.12) for u, we obtain

$$X_T^* = \frac{x}{C}\,(Z_T)^{-(1-q)}.$$

The optimal wealth process X^* is a $(Q, \mathbb{F})$-martingale, so

$$X_t^* = E^Q\big[X_T^* | \mathcal{F}_t\big] = \frac{1}{Z_t} E\big[Z_T X_T^* \,|\, \mathcal{F}_t\big] = \frac{x}{C Z_t} E\big[Z_T^q | \mathcal{F}_t\big], \quad 0 \le t \le T. \tag{3.21}$$

So, to compute explicit formulae for $C = E\big[Z_T^q \,|\, \mathcal{F}_0\big]$ and the optimal wealth process (from which the optimal trading strategy will be derived), we need to evaluate the conditional expectation $E\big[Z_T^q | \mathcal{F}_t\big], 0 \le t \le T$.

From (3.8), for $t \le T$, and conditional on $\mathcal{F}_t$, λ_T is normally distributed according to

$$\mathrm{Law}(\lambda_T | \mathcal{F}_t) = \mathrm{N}(\lambda_t, \mathrm{w}_t - \mathrm{w}_T), \quad 0 \le t \le T.$$

For a normally distributed random variable $Y \sim \mathrm{N}(m, s^2)$, we have

$$E \exp(cY^2) = \frac{1}{\sqrt{1-2cs^2}} \exp\left(\frac{cm^2}{1-2cs^2}\right),$$

so that, given the explicit expression (3.18) for Z_t, both C and the right-hand side of (3.21) can be computed in closed form. We find that C is indeed given by (3.13). Notice that $1 + q\mathrm{w}_0 T > 0$ and $1 + \mathrm{w}_0 T > 0$ due to the conditions on $\mathrm{w}_0 T$, thus the solution is well defined.

For the optimal wealth process, we obtain the formula

$$X_t^* = x\left(\frac{\Psi_t}{\Psi_0}\right)^{1/2} \exp\left(\frac{1}{2}(1-q)(\Phi_t - \Phi_0)\right), \quad 0 \le t \le T, \tag{3.22}$$

where

$$\Psi_t := \frac{\mathrm{w}_t}{1 + q\mathrm{w}_t(T-t)}, \quad \Phi_t := \frac{\lambda_t^2}{\mathrm{w}_t(1 + q\mathrm{w}_t(T-t))}, \quad 0 \le t \le T.$$

To compute the optimal trading strategy θ^*, we apply the Itô formula to (3.22), using the SDE for λ and noting that the derivative of w_t is given by

$$\frac{\mathrm{d}\mathrm{w}_t}{\mathrm{d}t} = -\mathrm{w}_t^2.$$

We compare the coefficient of $\mathrm{d}B_t$ in $\mathrm{d}X_t^*$ with that in (3.9) for the case of the optimal wealth process. This gives (3.14). □

3.2.1 Classical Merton problem

In the limit $\mathrm{w}_0 \to 0$, the drift of the stock becomes the constant λ_0, and Theorem 3.2 gives the solution to the classical full information Merton optimal investment problem for a stock with constant market price of risk λ_0 and volatility σ. In this case it is easy to check that the value function (3.12) becomes

$$u(x) = \frac{x^\gamma}{\gamma} \exp\left(\frac{1}{2}\frac{\gamma}{1-\gamma}\lambda_0^2 T\right),$$

and the optimal trading strategy (3.14) becomes

$$\theta_t^* = \frac{\lambda_0}{\sigma(1-\gamma)}, \quad 0 \le t \le T.$$

That is, the Merton investor keeps a constant proportion of wealth invested in the stock, as is well known.

3.3 Merton problem with uncertain drift

We can now solve the Merton problem when the agent has uncertainty over the true value of the drift parameter. Optimal investment models under partial information have been considered by many authors. We refer the reader to Rogers [31], Björk, Davis and Landén [1], and Platen and Runggaldier [30], for example.

A stock price process $S = (S_t)_{0 \le t \le T}$ follows

$$dS_t = \sigma S_t(\lambda dt + dB_t), \tag{3.23}$$

on a complete probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F} := (\mathcal{F}_t)_{0 \le t \le T}$, with $B = (B_t)_{0 \le t \le T}$ an $\mathbb{F}$-Brownian motion.

Define the process $\xi = (\xi_t)_{0 \le t \le T}$, by

$$\xi_t := \frac{1}{\sigma} \int_0^t \frac{dS_u}{S_u} = \lambda t + B_t. \tag{3.24}$$

The process ξ will be considered as the observation process in a filtering framework, corresponding to noisy observations of λ, with B representing the noise. In a partial information model with continuous stock price observations, an agent must use $\widehat{\mathbb{F}}$-adapted trading strategies, where where $\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_t)_{0 \le t \le T}$ is the observation filtration, defined by

$$\widehat{\mathcal{F}}_t := \sigma(\xi_s; 0 \le s \le t) = \sigma(S_s; 0 \le s \le t).$$

Then σ is known from the quadratic variation of S, but λ is an unknown constant, and hence modelled as an $\mathcal{F}_0$-measurable random variable. We assume the distribution of λ is Gaussian, $\lambda \sim \mathrm{N}(\lambda_0, \mathrm{v}_0)$, independent of B.

We are faced with a Kalman–Bucy type filtering problem whose unobservable signal process is the market price of risk λ. The signal process SDE is

$$d\lambda = 0, \tag{3.25}$$

and the observation process SDE is (3.24).

We apply Theorem 2.1 to the signal process λ in (3.25) and observation process ξ in (3.24). Then the optimal filter

$$\widehat{\lambda}_t := E\big[\lambda \mid \widehat{\mathcal{F}}_t\big], \quad 0 \le t \le T,$$

satisfies

$$d\widehat{\lambda}_t = \mathrm{v}_t d\widehat{B}_t, \quad \widehat{\lambda}_0 = \lambda_0, \tag{3.26}$$

where

$$\mathrm{v}_t := E\big[(\lambda - \widehat{\lambda}_t)^2 | \widehat{\mathcal{F}}_t\big], \quad 0 \le t \le T,$$

is the conditional variance of λ. This satisfies the Riccati equation

$$\frac{d\mathrm{v}_t}{dt} = -\mathrm{v}_t^2, \tag{3.27}$$

with initial value v_0, so that

$$\mathrm{v}_t = \frac{\mathrm{v}_0}{1 + \mathrm{v}_0 t}, \quad 0 \le t \le T. \tag{3.28}$$

The process $\widehat{B}$ is an $\widehat{\mathbb{F}}$-Brownian motion, the innovations process, satisfying

$$d\widehat{B}_t = d\xi_t - \widehat{\lambda}_t dt. \tag{3.29}$$

Using this in (3.26), the optimal filter can also be written in terms of the observable ξ as

$$\widehat{\lambda}_t = \frac{\lambda_0 + \mathrm{v}_0 \xi_t}{1 + \mathrm{v}_0 t}, \quad 0 \le t \le T. \tag{3.30}$$

The effect of the filtering is that the agent is now investing in a stock with dynamics given by $\mathrm{d}S_t = \sigma S_t \mathrm{d}\xi_t$ which, using (3.29), becomes

$$\mathrm{d}S_t = \sigma S_t (\widehat{\lambda}_t \mathrm{d}t + \mathrm{d}\widehat{B}_t). \tag{3.31}$$

Our agent has a power utility function (3.11) and may invest a portion of his wealth in shares and the remaining wealth in a cash account with zero interest rate (for simplicity). The ($\widehat{\mathbb{F}}$-adapted) wealth process X^0 then follows

$$\mathrm{d}X^0_t = \sigma \theta^0_t X^0_t (\widehat{\lambda}_t \mathrm{d}t + \mathrm{d}\widehat{B}_t), \quad X^0_0 = x, \tag{3.32}$$

where θ^0_t is the proportion of wealth invested in shares at time $t \in [0, T]$, an $\widehat{\mathbb{F}}$-adapted process satisfying $\int_0^T \left(\theta^0_t\right)^2 \mathrm{d}t < \infty$ almost surely, and such that $X^0_t \ge 0$ almost surely for all $t \in [0, T]$. Denote by $\mathcal{A}_0$ the set of such admissible strategies.

The objective is to maximise expected utility of terminal wealth over the $\widehat{\mathbb{F}}$-adapted admissible strategies. The value function is

$$u_0(x) := \sup_{\theta \in \mathcal{A}_0} E\big[U(X^0_T) \,|\, \widehat{\mathcal{F}}_0\big].$$

This may now be treated as a full information problem, with state dynamics given by (3.32).

We see from equations (3.26), (3.28) and (3.31), that the solution to the partial information optimal portfolio problem is given by Theorem 3.2, when we replace the process λ of Theorem 3.2 by $\widehat{\lambda}$, and replace $(\mathrm{w}_t)_{0 \le t \le T}$ by $(\mathrm{v}_t)_{0 \le t \le T}$. We have therefore proved the following result.

Theorem 3.3 (Merton problem with uncertain drift). *In a complete market with stock price process S given by (3.23), suppose an agent is restricted to using stock price adapted strategies to maximise expected utility of terminal wealth, with power utility function given by (3.11). Suppose further that the agent's prior distribution for λ is Gaussian, according to*

$$\mathrm{Law}(\lambda \,|\, \widehat{\mathcal{F}}_0) = \mathrm{N}(\lambda_0, \mathrm{v}_0),$$

and assume that

$$-1 < \mathrm{v}_0 T < \frac{1-\gamma}{\gamma}.$$

Then the agent's value function is given by

$$u_0(x) = \frac{x^\gamma}{\gamma} C_0^{1-\gamma}.$$

where

$$C_0 = \left[\frac{(1 + \mathrm{v}_0 T)^q}{1 + q\mathrm{v}_0 T}\right]^{1/2} \exp\left[-\frac{1}{2} \frac{q(1-q)\lambda_0^2 T}{1 + q\mathrm{v}_0 T}\right], \quad q = -\frac{\gamma}{1-\gamma}.$$

The optimal trading strategy is $\theta^{0,*} = (\theta_t^{0,*})_{0\le t\le T}$*, given by*

$$\theta_t^{0,*} = \frac{\widehat{\lambda}_t}{\sigma(1-\gamma)}\left(\frac{1}{1+q\mathrm{v}_t(T-t)}\right), \quad 0 \le t \le T,$$

where $\widehat{\lambda} = (\widehat{\lambda}_t)_{0\le t\le T}$ *satisfies (3.26) and* v_t *is given by (3.28).*

The classical Merton strategy is thus altered in two ways: the constant λ is replaced by its filtered estimate $\widehat{\lambda}_t$, and the risky asset proportion is decreased by the factor $(1 + q\mathrm{v}_t(T-t))^{-1}$. We note that the more risk averse the investor, the less likely he is to invest in shares, and as $t \to T$, the optimal strategy gets closer and closer to the Merton rule.

4 Investment with inside information and drift uncertainty

We again consider the Merton optimal investment problem in which the agent does not know the stock price drift, but now with the added feature that the agent has some additional information at time zero, represented by noisy knowledge of the terminal value B_T of the Brownian motion driving the stock. We refer the reader to Danilova, Monoyios and Ng [2] for further examples, such as when the additional information involves noisy knowledge of the terminal stock price. The work in this section and in [2] extends the classical inside information model of Pikovsky and Karatzas [29] by considering the situation where the insider does not know the stock's appreciation rate. The agent must use strategies that are adapted to the stock price filtration, but enlarged by the additional information. We must therefore utilise a filtering algorithm which computes the best estimate of the drift, given stock price observations and the additional information. The usual Kalman–Bucy equations hold in this scenario, but with modified initial conditions reflecting the additional information.

The market is the same one as in Section 3.3, with a single stock whose price process S follows (3.23), on a complete probability space $(\Omega, \mathcal{F}, P)$ equipped with a background filtration $\mathbb{F} := (\mathcal{F}_t)_{0\le t\le T}$, with B an $\mathbb{F}$-Brownian motion. We shall again allow for uncertainty in the value of λ, so consider it to be an $\mathcal{F}_0$-measurable random variable. Once again we take the interest rate to be zero.

As before, we define the observation process $\xi = (\xi_t)_{0\le t\le T}$ by (3.24), and the filtration generated by ξ is again denoted by $\widehat{\mathbb{F}} = (\widehat{\mathcal{F}}_t)_{0\le t\le T}$. Since the background filtration $\mathbb{F}$ contains the Brownian filtration and also the sigma-field generated by λ, we have $\widehat{\mathcal{F}}_t \subseteq \mathcal{F}_t$, for all $t \in [0, T]$.

Also as before, the uncertainty in the $\mathcal{F}_0$-measurable random variable λ is modelled by assuming that its prior distribution conditional on $\widehat{\mathcal{F}}_0$ is Gaussian, according to

$$\mathrm{Law}(\lambda \mid \widehat{\mathcal{F}}_0) = \mathrm{N}(\lambda_0, \mathrm{v}_0), \quad \text{independent of } B, \tag{4.1}$$

for given constants λ_0, v_0.[1]

In contrast to earlier, the utility-maximising agent will not only have access to $\widehat{\mathbb{F}}$ in order to estimate λ and implement an optimal strategy, but will be able to augment $\widehat{\mathbb{F}}$ with some additional information, represented by knowledge of a random variable I.

Our procedure in this section is to first enlarge the background filtration $\mathbb{F}$ with the information carried by the random variable I. Denote the enlarged filtration by $\mathbb{F}^{\sigma(I)} = (\mathcal{F}_t^{\sigma(I)})_{0\le t\le T}$, with

$$\mathcal{F}_t^{\sigma(I)} := \mathcal{F}_t \vee \sigma(I), \quad 0 \le t \le T.$$

By starting with an enlarged background filtration and then considering the optimal investment problem with uncertain drift, we aim to incorporate the insider's additional information in the estimation of the unknown market price of risk λ.

The next step is to write the stock price SDE (3.23) in terms of quantities adapted to $\mathbb{F}^{\sigma(I)}$. As $\mathbb{F}$ contains the Brownian filtration, we apply classical initial enlargement results (see, for instance, Mansuy and Yor [18]). There exists an $\mathbb{F}^{\sigma(I)}$-adapted process ν, the *information drift*, such that the Brownian motion B decomposes according to

$$B_t := B_t^I + \int_0^t \nu_s \mathrm{d}s, \quad 0 \le t \le T, \tag{4.2}$$

where B^I is an $\mathbb{F}^{\sigma(I)}$-Brownian motion. We shall characterise the information drift via Lemma 4.2 shortly.

Using (4.2), the stock price dynamics (3.23) is written in terms of $\mathbb{F}^{\sigma(I)}$-adapted processes, to give

$$\mathrm{d}S_t = \sigma S_t \left(\lambda_t^I \mathrm{d}t + \mathrm{d}B_t^I\right), \tag{4.3}$$

where

$$\lambda_t^I := \lambda + \nu_t, \quad 0 \le t \le T,$$

is $\mathbb{F}^{\sigma(I)}$-adapted. If the insider happened to know the value of λ, then we would interpret (4.3) as his stock price SDE, with a stochastic market price of risk λ^I, on the filtered probability space $(\Omega, \mathcal{F}_T^{\sigma(I)}, \mathbb{F}^{\sigma(I)}, P)$.

We study a problem where the inside information consists of noisy Brownian inside information. In other words, we take I to be given by

$$I := aB_T + (1-a)\epsilon, \quad 0 < a < 1, \tag{4.4}$$

and where ϵ is a standard normal random variable independent of B and λ.

Define the insider's observation filtration $\widehat{\mathbb{F}}^{\sigma(I)} = (\widehat{\mathcal{F}}_t^{\sigma(I)})_{0\le t\le T}$ by

$$\widehat{\mathcal{F}}_t^{\sigma(I)} := \sigma(I, \xi_s; 0 \le s \le t), \quad 0 \le t \le T.$$

We now incorporate the insider's uncertainty in the knowledge of λ by treating it as an $\mathcal{F}_0^{\sigma(I)}$-measurable Gaussian random variable with distribution conditional on $\widehat{\mathcal{F}}_0$ given

[1] One way to choose λ_0, v_0 would be to use past data before time zero to obtain a point estimate of λ, and to use the distribution of the estimator as the prior, as in Monoyios [23] and Section 5 of this article.

by (4.1). In this example, λ is independent of I, so its distribution conditional on $\widehat{\mathcal{F}}_0^{\sigma(I)}$ is unaltered from that in (4.1):

$$\mathrm{Law}(\lambda \mid \widehat{\mathcal{F}}_0^{\sigma(I)}) = \mathrm{Law}(\lambda \mid \widehat{\mathcal{F}}_0) = \mathrm{N}(\lambda_0, \mathrm{v}_0). \tag{4.5}$$

Treating λ^I as an unobservable signal process, we shall see that λ^I will satisfy a linear SDE with respect to $\mathbb{F}^{\sigma(I)}$. The Kalman–Bucy filter then allows the insider to infer the conditional expectation

$$\widehat{\lambda}_t^I := E\big[\lambda_t^I \mid \widehat{\mathcal{F}}_t^{\sigma(I)}\big], \quad 0 \le t \le T, \tag{4.6}$$

that is, the best estimate of the signal λ^I based on the insider's observation filtration $\widehat{\mathbb{F}}^{\sigma(I)}$, which turns out to be a Gaussian process, fully characterised by the filtering algorithm. The initial condition for the optimal filter incorporates the inside information, and the SDE for the filter augments this with the stock price observations. This will convert the partial information model (4.3) to a full information model on the filtered probability space $(\Omega, \widehat{\mathcal{F}}_T^{\sigma(I)}, \widehat{\mathbb{F}}^{\sigma(I)}, P)$ with the stock price following

$$\mathrm{d}S_t = \sigma S_t(\widehat{\lambda}_t^I \mathrm{d}t + \mathrm{d}\widehat{B}_t^I), \tag{4.7}$$

where $\widehat{B}^I$ is an $\widehat{\mathbb{F}}^{\sigma(I)}$-Brownian motion. Finally, once we have the full information model (4.7), we are able to compute the maximum utility via duality.

Denote the agent's $\widehat{\mathbb{F}}^{\sigma(I)}$-adapted wealth process by $X^I = (X_t^I)_{0\le t\le T}$, with trading strategy $\theta^I = (\theta_t^I)_{0\le t\le T}$, the proportion of wealth invested in the stock, an $\widehat{\mathbb{F}}^{\sigma(I)}$-adapted process satisfying $\int_0^T \left(\theta_t^I\right)^2 \mathrm{d}t < \infty$ almost surely, such that $X_t^I \ge 0$ almost surely for all $t \in [0,T]$. Denote by $\mathcal{A}_I$ the set of such admissible strategies.

The value function for this problem is

$$u_I(x) := \sup_{\theta^I \in \mathcal{A}_I} E\big[U(X_T^I) \mid \widehat{\mathcal{F}}_0^{\sigma(I)}\big], \quad x > 0, \tag{4.8}$$

where U is the power utility function (3.11). We emphasise that the objective function in (4.8) is conditioned on $\widehat{\mathcal{F}}_0^{\sigma(I)}$.

Define the modulated terminal time T_a by

$$T_a := T + \Big(\frac{1-a}{a}\Big)^2, \tag{4.9}$$

which will appear in our results. Then the solution to this problem is as follows.

Theorem 4.1. *Assume that*

$$\frac{T}{T_a} - 1 < \mathrm{v}_0 T < \frac{T}{T_a} + \frac{1-\gamma}{\gamma}.$$

Define the function $\mathrm{v}^I : [0,T] \to \mathbb{R}$ *by*

$$\mathrm{v}_t^I := \frac{\mathrm{v}_0^I}{1 + \mathrm{v}_0^I t}, \quad \mathrm{v}_0^I := \mathrm{v}_0 - \frac{1}{T_a}, \quad 0 \le t \le T. \tag{4.10}$$

Then the process $\widehat{\lambda}^I$ *in (4.6) is given by*

$$\widehat{\lambda}^I_t = \lambda_0 + \frac{I}{aT_a} + \int_0^t \mathrm{v}^I_s \mathrm{d}\widehat{B}^I_s, \quad 0 \le t \le T, \tag{4.11}$$

where I *is defined in (4.4) and* T_a *in (4.9). The value function of the insider with knowledge of* I *at time zero is given by*

$$u_I(x) = \frac{x^\gamma}{\gamma} C_I^{1-\gamma}, \tag{4.12}$$

where C_I *is the* $\widehat{\mathcal{F}}^I_0$*-measurable random variable given by*

$$C_I = \left(\frac{(1+\mathrm{v}^I_0 T)^q}{1+q\mathrm{v}^I_0 T}\right)^{1/2} \exp\left(-\frac{1}{2}\frac{q(1-q)(\widehat{\lambda}^I_0)^2 T}{1+q\mathrm{v}^I_0 T}\right), \quad q = -\frac{\gamma}{1-\gamma}.$$

The insider's optimal trading strategy is $\theta^{I,*} = (\theta^{I,*}_t)_{0\le t\le T}$*, given by*

$$\theta^{I,*}_t = \frac{\widehat{\lambda}^I_t}{\sigma(1-\gamma)}\left(\frac{1}{1+q\mathrm{v}^I_t(T-t)}\right), \quad 0 \le t \le T.$$

Of course, the value function (4.12) depends explicitly on I, through its dependence on $\widehat{\lambda}^I_0$. We note the similarity in the structure of the solution to this problem with that of the Merton problem with uncertain drift and no inside information. The function v^I plays a similar role to the function v in the conventional partial information problem. It turns out that v^I is related to (but not identical to) the variance of λ^I conditional on $\widehat{\mathbb{F}}^I$, as we shall see.

4.1 Computing the information drift

The first result we need in order to prove Theorem 4.1 is a lemma that gives an explicit formula for the information drift in (4.2). Recall that we begin with a background filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ that includes the Brownian filtration and the sigma-field generated by λ. We enlarge $\mathbb{F}$ with the information carried by the random variable I. Define, for a bounded Borel function $f : \mathbb{R} \to \mathbb{R}$, the process $(\pi_t(f))_{0\le t\le T}$ as the continuous version of the martingale $(E[f(I) \,|\, \mathcal{F}_t])_{0\le t\le T}$:

$$\pi_t(f) := E[f(I) \,|\, \mathcal{F}_t], \quad 0 \le t \le T.$$

There then exists a predictable family of measures $(\mu_t(\mathrm{d}x))_{0\le t\le T}$ such that

$$\pi_t(f) = \int_{\mathbb{R}} f(x)\mu_t(\mathrm{d}x).$$

For fixed $t \in [0,T]$, the measure $\mu_t(\mathrm{d}x)$ is the conditional distribution of I given $\mathcal{F}_t$. Suppose I is such that there exists a density function $g(t,x,y)$ for each $t \in [0,T]$, and such that

$$\pi_t(f) = \int_{\mathbb{R}} f(x)\mu_t(\mathrm{d}x) = \int_{\mathbb{R}} f(x) g(t,x,B_t)\mathrm{d}x. \tag{4.13}$$

The enlargement decomposition formula is given by the following lemma.

Lemma 4.2. *Suppose that I is continuous random variable with conditional (on $\mathcal{F}_t$) distribution given by $g(t, x, B_t)$. Assume also that this distribution satisfies the following conditions:*

$$\int_{\mathbb{R}} |g_y(t, x, y)| \, \mathrm{d}x < \infty, \qquad \int_{\mathbb{R}} \left| \frac{g_y(t, x, y)}{g(t, x, y)} \right| \mathrm{d}x < \infty,$$

for a.e. $t \in [0, T]$ and a.e. $y \in \mathbb{R}$. Then the $\mathbb{F}$-Brownian motion B decomposes with respect to the enlarged filtration $\mathbb{F}^{\sigma(I)}$ according to

$$B_t = B_t^I + \int_0^t \nu_s \mathrm{d}s, \quad 0 \le t \le T,$$

where B^I is an $\mathbb{F}^{\sigma(I)}$-Brownian motion. The information drift ν is given by

$$\nu_t = \frac{g_y(t, I, B_t)}{g(t, I, B_t)}, \quad 0 \le t \le T.$$

Proof. Let f be a test function. Introduce the $\mathbb{F}$-predictable process $(\dot{\pi}_t(f))_{0 \le t \le T}$ such that

$$\pi_t(f) = Ef(I) + \int_0^t \dot{\pi}_s(f) \mathrm{d}B_s,$$

which exists by the representation property of Brownian martingales as stochastic integrals with respect to B. There exists a predictable family of measures $(\dot{\mu}_t(\mathrm{d}x))_{0 \le t \le T}$ such that

$$\dot{\pi}_t(f) = \int_{\mathbb{R}} f(x) \dot{\mu}_t(\mathrm{d}x),$$

and such that for each $t \in [0, T]$ the measure $\dot{\mu}_t(\mathrm{d}x)$ is absolutely continuous with respect to $\mu_t(\mathrm{d}x)$. Define $\alpha(t, x)$ by

$$\dot{\mu}_t(\mathrm{d}x) = \alpha(t, x) \mu_t(\mathrm{d}x).$$

Now suppose we have a continuous $\mathbb{F}$-martingale M given by

$$M_t = \int_0^t m_s \mathrm{d}B_s, \qquad 0 \le t \le T.$$

By Theorem 1.6 in Mansuy and Yor [18], there exists an $\mathbb{F}^{\sigma(I)}$-local martingale M^I such that

$$M_t = M_t^I + \int_0^t \alpha(s, I) \, \mathrm{d}[M, B]_s,$$

provided that, almost surely,

$$\int_0^t |\alpha(s, I)| \, \mathrm{d}[M, B]_s < \infty.$$

In particular, if $\int_0^t |\alpha(s, I)| \, \mathrm{d}s < \infty$ almost surely, then B decomposes as

$$B_t = B_t^I + \int_0^t \alpha(s, I) \, \mathrm{d}s, \quad 0 \le t \le T,$$

with B^I an $\mathbb{F}^{\sigma(I)}$-Brownian motion.

From the definition of $\alpha(t,x)$ we have

$$\dot{\pi}_t(f) = \int_{\mathbb{R}} f(x)\alpha(t,x)\mu_t(\mathrm{d}x) = \int_{\mathbb{R}} f(x)\alpha(t,x)g(t,x,B_t)\mathrm{d}x.$$

Hence,

$$\mathrm{d}\pi_t(f) = \dot{\pi}_t(f)\mathrm{d}B_t = \left(\int_{\mathbb{R}} f(x)\alpha(t,x)g(t,x,B_t)\mathrm{d}x\right)\mathrm{d}B_t,$$

so that

$$\mathrm{d}[\pi(f),M]_t = \left(\int_{\mathbb{R}} f(x)\alpha(t,x)g(t,x,B_t)\mathrm{d}x\right)\mathrm{d}[B,M]_t. \tag{4.14}$$

But from the defining representation (4.13), the right-hand side of which is a smooth function of B_t, the Itô formula gives

$$\mathrm{d}[\pi(f),M]_t = \left(\int_{\mathbb{R}} f(x)g_y(t,x,B_t)dx\right)\mathrm{d}[B,M]_t, \tag{4.15}$$

and comparing (4.14) with (4.15) yields the result. □

Proof of Theorem 4.1. For I given by (4.4), the conditional distribution of I given $\mathcal{F}_t$, for $t \le T$, is

$$\mathrm{N}(aB_t, a^2(T-t) + (1-a)^2) = \mathrm{N}(aB_t, a^2(T_a - t)),$$

where T_a is defined in (4.9). Hence the conditional density is

$$g(t,x,B_t) = \frac{1}{a\sqrt{2\pi(T_a-t)}}\exp\left[-\frac{1}{2}\frac{(x-aB_t)^2}{a^2(T_a-t)}\right].$$

So by Lemma 4.2, the information drift is

$$\nu_t = \frac{I - aB_t}{a(T_a-t)}, \quad 0 \le t \le T. \tag{4.16}$$

Using the information drift in (4.16) we write the stock price SDE (3.23) in terms of $\mathbb{F}^{\sigma(I)}$-adapted processes, to obtain (4.3), where the $\mathbb{F}^{\sigma(I)}$-adapted market price of risk λ^I is given by

$$\lambda_t^I := \lambda + \nu_t = \lambda + \frac{I - aB_t}{a(T_a-t)} =: h(t,B_t), \quad 0 \le t \le T,$$

and where $h : [0,T] \times \mathbb{R} \to \mathbb{R}$ is defined by

$$h(t,x) := \lambda + \frac{I - ax}{a(T_a-t)}.$$

Applying the Itô's formula and using $\mathrm{d}B_t = \nu_t \mathrm{d}t + \mathrm{d}B_t^I$, we obtain

$$\mathrm{d}\lambda_t^I = -\frac{1}{T_a-t}\mathrm{d}B_t^I, \quad \lambda_0^I = \lambda + \frac{I}{aT_a}. \tag{4.17}$$

With ξ being the returns process in (3.24), we have

$$\mathrm{d}\xi_t = \lambda_t^I \mathrm{d}t + \mathrm{d}B_t^I. \tag{4.18}$$

We now regard λ as an unknown constant, and hence a random variable, whose distribution conditional on $\widehat{\mathcal{F}}_0^{\sigma(I)}$ is given by (4.5). Then we regard $(\lambda_t^I)_{0\le t\le T}$ as an unobservable signal process following (4.17), and ξ as an observation process following (4.18), in a filtering framework to estimate of λ_t^I conditional on $\widehat{\mathcal{F}}_t^{\sigma(I)}$.

Using (4.5), we can write down the initial distribution of λ_0^I given $\widehat{\mathcal{F}}_0^{\sigma(I)}$:

$$\mathrm{Law}(\lambda_0^I|\widehat{\mathcal{F}}_0^{\sigma(I)}) = \mathrm{Law}\left(\lambda + \frac{I}{aT_a}\middle|\,\widehat{\mathcal{F}}_0^{\sigma(I)}\right) = \mathrm{N}\left(\lambda_0 + \frac{I}{aT_a}, \mathrm{v}_0\right).$$

This defines the prior distribution of the signal process λ^I. Of course, since I is $\widehat{\mathcal{F}}_0^{\sigma(I)}$-measurable, it does not contribute to the initial variance.

The Kalman–Bucy filter, Theorem 2.1, is directly applicable, and yields that the optimal filter

$$\widehat{\lambda}_t^I := E\big[\lambda_t^I \mid \widehat{\mathcal{F}}_t^{\sigma(I)}\big], \quad 0 \le t \le T,$$

satisfies the SDE

$$\mathrm{d}\widehat{\lambda}_t^I = \left(V_t^I - \frac{1}{T_a - t}\right)\mathrm{d}\widehat{B}_t^I, \quad \widehat{\lambda}_0^I = \lambda_0 + \frac{I}{aT_a}, \tag{4.19}$$

where $\widehat{B}^I$ is the innovations process, an $\widehat{\mathbb{F}}^{\sigma(I)}$-Brownian motion defined by

$$\widehat{B}_t^I := \xi_t - \int_0^t \widehat{\lambda}_s^I \mathrm{d}s, \quad 0 \le t \le T, \tag{4.20}$$

and V_t^I is the conditional variance of λ_t^I:

$$V_t^I := E\Big[\,(\lambda_t^I - \widehat{\lambda}_t^I)^2\Big|\,\widehat{\mathcal{F}}_t^{\sigma(I)}\Big], \quad 0 \le t \le T,$$

which satisfies

$$\frac{\mathrm{d}V_t^I}{\mathrm{d}t} = \frac{2}{T_a - t}V_t^I - \left(V_t^I\right)^2, \quad V_0^I = \mathrm{v}_0.$$

If we define

$$\mathrm{v}_t^I := V_t^I - \frac{1}{T_a - t}, \quad 0 \le t \le T,$$

then (4.19) becomes

$$\mathrm{d}\widehat{\lambda}_t^I = \mathrm{v}_t^I \mathrm{d}\widehat{B}_t^I, \quad \widehat{\lambda}_0^I = \lambda_0 + \frac{I}{aT_a}. \tag{4.21}$$

Note that (4.21) is of the same form as (3.8) with w_t replaced by v_t^I and with B_t replaced by $\widehat{B}_t^I$. Indeed, v_t^I plays the role of an 'effective variance', satisfying the Riccati equation (3.27), with a modified initial condition:

$$\frac{\mathrm{d}\mathrm{v}_t^I}{\mathrm{d}t} = -\left(\mathrm{v}_t^I\right)^2, \quad \mathrm{v}_0^I = \mathrm{v}_0 - \frac{1}{T_a}.$$

The solution to this equation is then given by (4.10), and the solution to (4.21) is then (4.11).

Using (4.20) in the SDE (4.21), the optimal filter may also be written explicitly in terms of the observable ξ, as

$$\widehat{\lambda}_t^I = \frac{\widehat{\lambda}_0^I + \mathrm{v}_0^I \xi_t}{1 + \mathrm{v}_0^I t}, \quad 0 \le t \le T.$$

This is of the same form as (3.30), with λ_0 replaced by $\widehat{\lambda}_0^I$ and v_0 replaced by v_0^I.

The effect of the filtering is that the agent is now investing in a stock with dynamics given by $\mathrm{d}S_t = \sigma S_t \mathrm{d}\xi_t$ which, using (4.20), becomes (4.7). The $\widehat{\mathbb{F}}^{\sigma(I)}$-adapted wealth process X^I then follows

$$\mathrm{d}X_t^I = \sigma \theta_t^I X_t^I (\widehat{\lambda}_t^I \mathrm{d}t + \mathrm{d}\widehat{B}_t^I), \quad X_0^I = x,$$

where θ^I is the $\widehat{\mathbb{F}}^{\sigma(I)}$-adapted trading strategy. The theorem then follows immediately from making the replacements

$$\mathrm{w} \to \mathrm{v}^I, \quad \lambda \to \widehat{\lambda}^I,$$

in Theorem 3.2. □

It can be shown that the additional information increases the insider's utility over the regular agent: see [2] for this and other effects of the inside information.

5 Optimal hedging of basis risk with partial information

In this section we analyse the hedging of a contingent claim in a basis risk model, a tractable example of an incomplete market, first under a full information assumption, and then under a partial information scenario. Basis risk models involve a claim on a non-traded asset, which is hedged using a correlated traded asset. They were first studied systematically by Davis [4] (whose preprint on the subject originated in 2000) who used a dual approach to derive approximations for indifference prices. Subsequently, Henderson [8], and Musiela and Zariphopoulou [26] derived an expectation representation (given in Theorem 5.3) for the value function of the utility maximisation problem involving a random endowment of the claim. This was used by Monoyios [21] to derive accurate analytic approximations for indifference prices and hedging strategies. In simulation experiments, Monoyios showed that exponential indifference hedging could outperform the BS approximation of taking the traded asset as a good proxy for the non-traded asset. Unfortunately, the utility-based hedge requires knowledge of the drift parameters of the assets. These are hard to estimate accurately, as shown by Rogers [31] and Monoyios [22], who showed that drift parameter mis-estimation could ruin the effectiveness of the optimal hedge. Finally, in [23, 25] Monoyios developed a filtering algorithm to deal with the drift parameter uncertainty, and showed that with this added ingredient, utility-based hedging was indeed effective, even in the face of parameter uncertainty. We shall describe some of these results in this section.

5.1 Basis risk model: full information case

In a full information model, the setting is a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{0 \le t \le T}, P)$, where the filtration $\mathbb{F}$ is the P-augmentation of that generated by a two-dimensional Brownian motion $(B, B^{\perp})$. A traded stock price $S := (S_t)_{0 \le t \le T}$ follows a log-Brownian process given by

$$\mathrm{d}S_t = \quad \sigma S_t(\lambda \mathrm{d}t + \mathrm{d}B_t) \quad =: \sigma S_t \mathrm{d}\xi_t, \tag{5.1}$$

where $\sigma > 0$ and λ are known constants. For simplicity, the interest rate is taken to be zero. The process ξ in (5.1) defined by $\mathrm{d}\xi_t := \lambda \mathrm{d}t + \mathrm{d}B_t$ will subsequently play a role as one component of an observation process in a partial information model, when λ will be treated as a random variable rather than as a known constant.

A non-traded asset price $Y := (Y_t)_{0 \le t \le T}$ follows the correlated log-Brownian motion

$$\mathrm{d}Y_t = \quad \beta Y_t(\theta \mathrm{d}t + \mathrm{d}W_t) \quad =: \beta Y_t \mathrm{d}\zeta_t, \tag{5.2}$$

with $\beta > 0$ and θ known constants. The Brownian motion W is correlated with B according to

$$[B, W]_t = \rho t, \quad W = \rho B + \sqrt{1 - \rho^2} B^{\perp}, \quad \rho \in [-1, 1],$$

and the process ζ, given by $\mathrm{d}\zeta_t := \theta \mathrm{d}t + \mathrm{d}W_t$, will act as the second component of an observation process in a partial information model, when θ will be considered a random variable. We shall henceforth refer to the Sharpe ratios λ (respectively, θ) as the drift of S (respectively, Y), for brevity.

A European contingent claim pays the non-negative random variable $h(Y_T)$ at time T, where $h : \mathbb{R}^+ \to \mathbb{R}^+$. In what follows we shall consider utility maximisation problems with the additional random terminal endowment $nh(Y_T)$, for $n \in \mathbb{R}$. We assume the random endowment $nh(Y_T)$ is continuous and bounded below, with finite expectation under any martingale measure.

An agent may trade the stock in a self-financing fashion, leading to the portfolio wealth process $X = (X_t)_{0 \le t \le T}$ satisfying

$$\mathrm{d}X_t = \sigma \pi_t(\lambda \mathrm{d}t + \mathrm{d}B_t),$$

where $\pi := (\pi_t)_{0 \le t \le T}$ is the wealth in the stock, representing the agent's trading strategy, satisfying $\int_0^T \pi_t^2 \mathrm{d}t < \infty$ almost surely.

5.1.1 Perfect correlation case

This market is incomplete for $|\rho| \neq 1$. If the correlation is perfect, however, the market becomes complete and perfect hedging is possible, as we now show.

The minimal martingale measure Q^M has density process with respect to P given by

$$\left.\frac{\mathrm{d}Q^M}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \mathcal{E}\left(-\lambda \cdot B\right)_t, \quad 0 \le t \le T.$$

Under Q^M, (S, Y) follow

$$\begin{aligned} \mathrm{d}S_t &= \sigma S_t \mathrm{d}B_t^{Q^M}, \\ \mathrm{d}Y_t &= \beta(\theta - \rho\lambda) Y_t \mathrm{d}t + \beta Y_t \mathrm{d}W_t^{Q^M}, \end{aligned} \tag{5.3}$$

where B^{Q^M}, W^{Q^M} are correlated Brownian motions under Q^M. The stock price S is a local Q^M-martingale, but this is not the case for the non-traded asset, unless we have the perfect correlation case, $\rho = 1$. In this case Y is effectively a traded asset (as Y_t is then a deterministic function of S_t), so the Q^M-drift of Y vanishes. Therefore, given σ, β, when $\rho = 1$ the Sharpe ratios λ, θ are equal:

$$\theta = \lambda.$$

In this case the market becomes complete, and perfect hedging is possible. It is easy to show that with $\rho = 1$, so that $W = B$, we have

$$Y_t = Y_0 \left(\frac{S_t}{S_0}\right)^{\beta/\sigma} e^{ct}, \quad c = \frac{1}{2}\sigma\beta\left(1 - \frac{\beta}{\sigma}\right).$$

Let the claim price process be $v(t, Y_t), 0 \leq t \leq T$, where $v : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+$ is smooth enough to apply the Itô formula, so that

$$\mathrm{d}v(t, Y_t) = \left[v_t(t, Y_t) + \mathcal{A}^Y v(t, Y_t)\right] \mathrm{d}t + \beta Y_t v_y(t, Y_t) \mathrm{d}W_t,$$

where $\mathcal{A}^Y$ is the generator of the process Y in (5.2). The replication conditions are

$$X_t = v(t, Y_t), \quad 0 \leq t \leq T, \quad \mathrm{d}X_t = \mathrm{d}v(t, Y_t).$$

Standard arguments then show that to perfectly hedge the claim one must hold Δ_t shares of S at $t \in [0, T]$, given by

$$\Delta_t = \frac{\beta}{\sigma}\frac{Y_t}{S_t}\frac{\partial v}{\partial y}(t, Y_t), \quad 0 \leq t \leq T, \tag{5.4}$$

and the claim pricing function $v(t, y)$ satisfies

$$v_t(t, y) + \beta(\theta - \lambda) y v_y(t, y) + \frac{1}{2}\beta^2 y^2 v_{yy}(t, y) = 0, \quad v(T, y) = h(y).$$

But with $\rho = 1$, $\theta = \lambda$, so we get the BS partial differential equation (PDE), and hence

$$v(t, Y_t) = \mathrm{BS}(t, Y_t), \quad 0 \leq t \leq T,$$

where $\mathrm{BS}(t, y)$ denotes the BS option pricing formula at time t, with underlying asset price y.

Therefore, a position in n claims is hedged by $\Delta_t^{(\mathrm{BS})}$ units of S at $t \in [0, T]$, where

$$\Delta_t^{(\mathrm{BS})} = -n\frac{\beta}{\sigma}\frac{Y_t}{S_t}\frac{\partial}{\partial y}\mathrm{BS}(t, Y_t; \beta), \quad 0 \leq t \leq T, \tag{5.5}$$

and where $\mathrm{BS}(t, y; \beta)$ denotes the BS formula at time t for underlying asset price y and volatility β. From our perspective, the salient feature of (5.5) is that the perfect hedge does not require knowledge of the values of the drifts λ, θ.

5.1.2 Incomplete case

Now suppose the correlation is not perfect, so that the market is incomplete. We embed the problem in a utility maximisation framework in a manner that is by now classical. Let the agent have risk preferences expressed via the exponential utility function

$$U(x) = -\exp(-\alpha x), \quad x \in \mathbb{R}, \quad \alpha > 0.$$

The agent maximises expected utility of terminal wealth at time T, with a random endowment of n units of claim payoff:

$$J(t, x, y; \pi) = E\big[U(X_T + nh(Y_T)) \mid X_t = x, Y_t = y\big].$$

The value function is $u^{(n)}(t, x, y) \equiv u(t, x, y)$, defined by

$$u(t, x, y) := \sup_{\pi \in \mathcal{A}} J(t, x, y; \pi), \quad u(T, x, y) = U(x + nh(y)), \tag{5.6}$$

where $\mathcal{A}$ is the set of admissible strategies. This is composed of S-integrable processes whose gains process is a Q-martingale for any martingale measure with finite relative entropy with respect to P. Denote the optimal trading strategy that achieves the supremum in (5.6) by $\pi^* \equiv \pi^{*,n}$, and denote the optimal wealth process by $X^* \equiv X^{*,n}$.

The following definitions of utility-based price and hedging strategy are now standard.

Definition 5.1 (Indifference price). The indifference price per claim at $t \in [0, T]$, given $X_t = x, Y_t = y$, is $p(t, x, y) \equiv p^{(n)}(t, x, y)$, defined by

$$u^{(n)}(t, x - np^{(n)}(t, x, y), y) = u^{(0)}(t, x, y).$$

We allow for possible dependence on t, x, y of $p^{(n)}$ in the above definition, but with exponential preferences it turns out that there is no dependence on x.

Definition 5.2 (Optimal hedging strategy). The optimal hedging strategy for n units of the claim is $\pi^H := (\pi_t^H)_{0 \le t \le T}$ given by

$$\pi_t^H := \pi_t^{*,n} - \pi_t^{*,0}, \quad 0 \le t \le T.$$

We have the following representation for the value function and indifference price.

Theorem 5.3. *The value function $u^{(n)}$ and indifference price $p^{(n)}$, given $X_t = x, Y_t = y$ for $t \in [0, T]$, are given by*

$$\begin{aligned} u^{(n)}(t, x, y) &= -e^{-\alpha x - \frac{1}{2}\lambda^2 (T-t)} \left[F(t, Y_t)\right]^{1/(1-\rho^2)}, \\ F(t, y) &= E^{Q^M}\left[\exp\left(-\alpha(1-\rho^2) nh(Y_T)\right) \big| \, Y_t = y\right], \qquad (5.7) \\ p^{(n)}(t, y) &= -\frac{1}{\alpha(1-\rho^2)n} \log F(t, y). \qquad (5.8) \end{aligned}$$

Proof. The Hamilton–Jacobi–Bellman (HJB) equation for the value function $u^{(n)}$ is

$$u_t^{(n)} + \sigma \sup_{\pi} \left(\lambda \pi u_x^{(n)} + \frac{1}{2}\sigma \pi^2 u_{xx}^{(n)} + \rho\beta\pi y u_{xy}^{(n)} \right) + \mathcal{A}^Y u^{(n)} = 0.$$

Performing the maximisation gives the optimal feedback control as $\Pi^{*,n}(t,x,y)$, where the function $\Pi^{*,n} : [0,T] \times \mathbb{R} \times \mathbb{R}^+$ is given by

$$\Pi^{*,n}(t,x,y) := -\left(\frac{\lambda u_x^{(n)} + \rho\beta y u_{xy}^{(n)}}{\sigma u_{xx}^{(n)}} \right). \tag{5.9}$$

The optimal trading strategy $\pi^{*,n}$ is then given by $\pi_t^{*,n} = \Pi^*(t, X_t^*, Y_t)$. Substituting the optimal Markov control back into the Bellman equation gives the HJB PDE

$$u_t^{(n)} + \mathcal{A}^Y u^{(n)} - \frac{\left(\lambda u_x^{(n)} + \rho\beta y u_{xy}^{(n)} \right)^2}{2u_{xx}^{(n)}} = 0.$$

The function $F(t,y)$ in (5.7) satisfies the linear PDE

$$F_t + \beta(\theta - \rho\lambda)F_y + \frac{1}{2}\beta^2 y^2 F_{yy} = 0, \quad F(T,y) = \exp(-\alpha(1-\rho^2)nh(y)),$$

by virtue of the Feynman–Kac theorem. It is then straightforward to verify that $u^{(n)}$ as given in the theorem solves the above HJB equation, and the definition of the indifference price gives the formula (5.8). □

This leads to the following representation for the optimal hedging strategy.

Theorem 5.4. *The optimal hedging strategy for a position in n claims is to hold Δ_t^H shares at $t \in [0,T]$, given by*

$$\Delta_t^H = -n\rho \frac{\beta}{\sigma} \frac{Y_t}{S_t} \frac{\partial p^{(n)}}{\partial y}(t, Y_t), \quad 0 \le t \le T. \tag{5.10}$$

Proof. From Theorem 5.3 the value function may be written in terms of the indifference price as

$$u^{(n)}(t,x,y) = -\exp\Big(-\alpha(x + np^{(n)}(t,y)) - \frac{1}{2}\lambda^2(T-t) \Big). \tag{5.11}$$

The optimal trading strategy is $\pi_t^{*,n} = \Pi^*(t, X_t^*, Y_t)$, where the function $\Pi^{*,n}(t,x,y)$ is given in (5.9), in terms of derivatives of the value function. Using (5.11) we obtain

$$\pi_t^{*,n} = \frac{\lambda - \rho\alpha n\beta Y_t p_y^{(n)}(t, Y_t)}{\alpha\sigma}, \quad 0 \le t \le T.$$

The optimal trading strategy for the problem with no claims, $\pi_t^{*,0}$ is obtained trivially by setting $n = 0$ in this result, and then applying Definition 5.2 proves the theorem. □

Notice that, given the PDE satisfied by F, the indifference pricing function $p(t,y) \equiv p^{(n)}(t,y)$ satisfies

$$p_t + \beta(\theta - \rho\lambda)yp_y + \frac{1}{2}\beta^2y^2p_{yy} - \frac{1}{2}\beta^2y^2\alpha n(1-\rho^2)(p_y)^2 = 0.$$

So for $\rho = 1$, in which case $\theta = \lambda$, the claim price then satisfies the BS PDE and we recover the perfect delta hedge (5.4).

In [21, 22] the hedging strategy in (5.10) is shown to be superior to the BS-style hedge (5.5), in terms of the terminal hedging error distribution produced by selling the claim at the appropriate price (the indifference price or the BS price) and investing the proceeds in the corresponding hedging portfolio. But from (5.3) we see that the exponential hedge requires knowledge of λ, θ, which are impossible to estimate accurately (see Rogers [31] or Monoyios [22]). This can ruin the effectiveness of indifference hedging, as shown in [22]. It is therefore dubious to draw any meaningful conclusions on the effectiveness of utility-based hedging in this model without relaxing the assumption that the agent knows the true values of the drifts.

5.2 Partial information case

Now we assume the hedger does not know the values of the return parameters λ, θ, so these are considered to be random variables. Equivalently, the agent cannot observe the Brownian motions B, W driving the asset prices, so is required to use strategies adapted to the observation filtration $\widehat{\mathbb{F}}$ generated by asset returns.

5.2.1 Choice of prior

We take the the two-dimensional random variable

$$\Xi := \begin{pmatrix} \lambda \\ \theta \end{pmatrix}$$

to have a Gaussian distribution which will be updated as the agent attempts to filter the values of the drifts from asset observations during the hedging interval $[0,T]$.

The choice of Gaussian prior is motivated by the idea that the agent has some past observations of S, Y before time zero, uses these to obtain classical point estimates of the drifts, and the joint distribution of the estimators is used as the prior in a Bayesian framework. Ultimately, in order to obtain explicit solutions, we shall assume that the agent uses observations before time zero of equal length for both assets. In setting the prior this way, we make the approximation that the asset price observations are continuous, so that σ, β, ρ are known from the quadratic variation and co-variation of S, Y. This is because our goal here is to focus on the severest problem of drift parameter uncertainty.

So, consider, for the moment, an observer with data for S over a time interval of length t_S, and for Y over a window of length t_Y, who considers λ and θ as *constants*, and records the returns $\mathrm{d}S_t/S_t$ and $\mathrm{d}Y_t/Y_t$ in order to estimate the values of the drifts.

An unbiased estimator of λ is $\bar{\lambda}(t_S)$ given by

$$\bar{\lambda}(t_S) = \frac{1}{t_S}\int_{t_0}^{t_0+t_S} \frac{\mathrm{d}S_u}{\sigma S_u} = \lambda + \frac{B_{t_0+t_S}}{t_S} \sim \mathrm{N}\left(\lambda, \frac{1}{t_S}\right).$$

The estimator of λ is normally distributed, with a similar computation for the estimator of θ. The estimator, $(\bar{\lambda}, \bar{\theta})$, of the (supposed constant) vector (λ, θ) is bivariate normal. Defining $\mathrm{v}_0 := 1/t_S$ and $\mathrm{w}_0 := 1/t_Y$ it is easily checked that

$$\begin{pmatrix} \bar{\lambda} \\ \bar{\theta} \end{pmatrix} \sim \mathrm{N}(M, C_0),$$

where the mean vector M and covariance matrix C_0 are given by

$$M = \begin{pmatrix} \lambda \\ \theta \end{pmatrix}, \qquad C_0 = \begin{pmatrix} \mathrm{v}_0 & \rho \min(\mathrm{v}_0, \mathrm{w}_0) \\ \rho \min(\mathrm{v}_0, \mathrm{w}_0) & \mathrm{w}_0 \end{pmatrix}. \tag{5.12}$$

With this in mind, we shall suppose that (λ, θ), now considered as a *random variable*, is bivariate normal according to

$$\lambda \sim \mathrm{N}(\lambda_0, \mathrm{v}_0), \quad \theta \sim \mathrm{N}(\theta_0, \mathrm{w}_0), \quad \mathrm{cov}(\lambda, \theta) = \mathrm{c}_0 := \rho \min(\mathrm{v}_0, \mathrm{w}_0),$$

for some chosen values λ_0, θ_0, typically obtained from past data prior to time zero. This distribution will be updated via subsequent observations of

$$\xi_t := \frac{1}{\sigma}\int_0^t \frac{\mathrm{d}S_u}{S_u} = \lambda t + B_t, \qquad \zeta_t := \frac{1}{\beta}\int_0^t \frac{\mathrm{d}Y_u}{Y_u} = \theta t + W_t,$$

over the hedging interval $[0, T]$.

5.2.2 Two-dimensional Kalman–Bucy filter

We are firmly within the realm of a two-dimensional Kalman filtering problem, which we treat as follows. Define the observation filtration by

$$\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_t)_{0 \le t \le T}, \qquad \widehat{\mathcal{F}}_t = \sigma(\xi_s, \zeta_s; 0 \le s \le t).$$

The observation process, Λ, and unobservable signal process, Ξ, are defined by

$$\Lambda := \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}_{0 \le t \le T}, \qquad \Xi := \begin{pmatrix} \lambda \\ \theta \end{pmatrix},$$

satisfying the stochastic differential equations

$$\mathrm{d}\Lambda_t = \Xi \mathrm{d}t + \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \mathrm{d}\begin{pmatrix} B_t \\ B_t^{\perp} \end{pmatrix}, \qquad \mathrm{d}\Xi = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The optimal filter is $\widehat{\Xi}_t := E[\Xi \mid \widehat{\mathcal{F}}_t], 0 \leq t \leq T$, a two-dimensional process defining the best estimates of λ and θ given observations up to time $t \in [0, T]$:

$$\widehat{\Xi}_t \equiv \begin{pmatrix} \widehat{\lambda}_t \\ \widehat{\theta}_t \end{pmatrix} := \begin{pmatrix} E[\lambda \mid \widehat{\mathcal{F}}_t] \\ E[\theta \mid \widehat{\mathcal{F}}_t] \end{pmatrix}, \qquad \begin{pmatrix} \widehat{\lambda}_0 \\ \widehat{\theta}_0 \end{pmatrix} = \begin{pmatrix} \lambda_0 \\ \theta_0 \end{pmatrix}. \tag{5.13}$$

The solution to this filtering problem converts the partial information model to a full information model with random drifts, given in the following proposition. To avoid a proliferation of symbols, we abuse notation and write $\widehat{\lambda}_t \equiv \widehat{\lambda}(t, S_t)$ and $\widehat{\theta} \equiv \widehat{\theta}(t, Y_t)$ for processes $\widehat{\lambda}, \widehat{\theta}$ that will turn out to be functions of time and current asset price.

Proposition 5.5. *The partial information model is equivalent to a full information model in which the asset price dynamics in the observation filtration $\widehat{\mathbb{F}}$ are*

$$\mathrm{d}S_t = \sigma S_t(\widehat{\lambda}_t \mathrm{d}t + \mathrm{d}\widehat{B}_t), \tag{5.14}$$

$$\mathrm{d}Y_t = \beta Y_t(\widehat{\theta}_t \mathrm{d}t + \mathrm{d}\widehat{W}_t), \tag{5.15}$$

where $\widehat{B}, \widehat{W}$ are $\widehat{\mathbb{F}}$-Brownian motions with correlation ρ, and the random drifts $\widehat{\lambda}, \widehat{\theta}$ are $\widehat{\mathbb{F}}$-adapted processes.

If λ and θ have common initial variance v_0, then $\widehat{\lambda}, \widehat{\theta}$ are given by

$$\begin{pmatrix} \widehat{\lambda}_t \\ \widehat{\theta}_t \end{pmatrix} = \begin{pmatrix} \lambda_0 \\ \theta_0 \end{pmatrix} + \int_0^t \mathrm{v}_s \begin{pmatrix} \mathrm{d}\widehat{B}_s \\ \mathrm{d}\widehat{W}_s \end{pmatrix}, \quad 0 \leq t \leq T, \tag{5.16}$$

where $(\mathrm{v}_t)_{0 \leq t \leq T}$ is the deterministic function

$$\mathrm{v}_t := \frac{\mathrm{v}_0}{1 + \mathrm{v}_0 t}, \quad 0 \leq t \leq T.$$

Equivalently, $\widehat{\lambda}, \widehat{\theta}$ are given as functions of time and current asset price by

$$\widehat{\lambda}_t = \widehat{\lambda}(t, S_t) = \frac{\lambda_0 + \mathrm{v}_0 \xi_t}{1 + \mathrm{v}_0 t}, \qquad \widehat{\theta}_t = \widehat{\theta}(t, Y_t) = \frac{\theta_0 + \mathrm{v}_0 \zeta_t}{1 + \mathrm{v}_0 t}, \tag{5.17}$$

with

$$\xi_t = \frac{1}{\sigma} \log\left(\frac{S_t}{S_0}\right) + \frac{1}{2}\sigma t, \qquad \zeta_t = \frac{1}{\beta} \log\left(\frac{Y_t}{Y_0}\right) + \frac{1}{2}\beta t. \tag{5.18}$$

Proof. Using a two-dimensional Kalman–Bucy filter (see, for example, Theorem V.9.2 in Fleming and Rishel [6]), $\widehat{\Xi}$ satisfies the stochastic differential equation

$$\mathrm{d}\widehat{\Xi}_t = C_t \left(DD^T\right)^{-1} (\mathrm{d}\Lambda_t - \widehat{\Xi}_t \mathrm{d}t) =: C_t \left(DD^T\right)^{-1} dN_t, \tag{5.19}$$

where $(N_t)_{0 \leq t \leq T}$ is the innovations process, defined by

$$N_t := \Lambda_t - \int_0^t \widehat{\Xi}_s \mathrm{d}s = \begin{pmatrix} \xi_t - \int_0^t \widehat{\lambda}_s \mathrm{d}s \\ \zeta_t - \int_0^t \widehat{\theta}_s \mathrm{d}s \end{pmatrix} =: \begin{pmatrix} \widehat{B}_t \\ \widehat{W}_t \end{pmatrix}, \tag{5.20}$$

and $\widehat{B}, \widehat{W}$ are $\widehat{\mathbb{F}}$-Brownian motions with correlation ρ. The deterministic matrix function C_t is the conditional variance-covariance matrix defined by

$$C_t := E\left[(\Xi - \widehat{\Xi}_t)(\Xi - \widehat{\Xi}_t)^{\mathrm{T}} \middle| \widehat{\mathcal{F}}_t\right] = E\left[(\Xi - \widehat{\Xi}_t)(\Xi - \widehat{\Xi}_t)^{\mathrm{T}}\right],$$

(T denoting transpose) where the last equality follows because the error $\Xi - \widehat{\Xi}_t$ is independent of $\widehat{\mathcal{F}}_t$ (Theorem V.9.2 in [6] again).

Using (5.20), and writing $\mathrm{d}S_t$ in terms of $\mathrm{d}\xi_t$, as in (5.1), gives the dynamics (5.14) of S in the observation filtration; (5.15) is established similarly.

The matrix $C = (C_t)_{0 \le t \le T}$ satisfies the Riccati equation

$$\frac{\mathrm{d}C_t}{\mathrm{d}t} = -C_t \left(DD^T\right)^{-1} C_t,$$

with C_0 given in (5.12). Then $R_t := C_t^{-1}$ satisfies the Lyapunov equation

$$\frac{\mathrm{d}R_t}{\mathrm{d}t} = \left(DD^T\right)^{-1}.$$

Define the elements of the conditional covariance matrix by

$$C_t =: \begin{pmatrix} \mathrm{v}_t & \mathrm{c}_t \\ \mathrm{c}_t & \mathrm{w}_t \end{pmatrix}.$$

Then the filtering equation (5.19) is a pair of coupled stochastic differential equations:

$$\begin{aligned} \begin{pmatrix} \mathrm{d}\widehat{\lambda}_t \\ \mathrm{d}\widehat{\theta}_t \end{pmatrix} &= \frac{1}{1-\rho^2} \begin{pmatrix} \mathrm{v}_t - \rho \mathrm{c}_t & \mathrm{c}_t - \rho \mathrm{v}_t \\ \mathrm{c}_t - \rho \mathrm{w}_t & \mathrm{w}_t - \rho \mathrm{c}_t \end{pmatrix} \begin{pmatrix} \mathrm{d}\xi_t - \widehat{\lambda}_t \mathrm{d}t \\ \mathrm{d}\zeta_t - \widehat{\theta}_t \mathrm{d}t \end{pmatrix} \\ &= \frac{1}{1-\rho^2} \begin{pmatrix} \mathrm{v}_t - \rho \mathrm{c}_t & \mathrm{c}_t - \rho \mathrm{v}_t \\ \mathrm{c}_t - \rho \mathrm{w}_t & \mathrm{w}_t - \rho \mathrm{c}_t \end{pmatrix} \begin{pmatrix} \mathrm{d}\widehat{B}_t \\ \mathrm{d}\widehat{W}_t \end{pmatrix}. \end{aligned}$$

Solving the Lyapunov equation yields three equations for $\mathrm{v}_t, \mathrm{w}_t, \mathrm{c}_t$:

$$\begin{aligned} \frac{\mathrm{v}_t}{\mathrm{v}_t \mathrm{w}_t - \mathrm{c}_t^2} - \frac{\mathrm{v}_0}{\mathrm{v}_0 \mathrm{w}_0 - \mathrm{c}_0^2} &= \frac{t}{1-\rho^2}, \\ \frac{\mathrm{w}_t}{\mathrm{v}_t \mathrm{w}_t - \mathrm{c}_t^2} - \frac{\mathrm{w}_0}{\mathrm{v}_0 \mathrm{w}_0 - \mathrm{c}_0^2} &= \frac{t}{1-\rho^2}, \\ \frac{\mathrm{c}_t}{\mathrm{v}_t \mathrm{w}_t - \mathrm{c}_t^2} - \frac{\mathrm{c}_0}{\mathrm{v}_0 \mathrm{w}_0 - \mathrm{c}_0^2} &= \frac{\rho t}{1-\rho^2}, \end{aligned} \tag{5.21}$$

where we have written $\mathrm{c}_0 \equiv \rho \min(\mathrm{v}_0, \mathrm{w}_0)$ for brevity.

Now make the simplification $\mathrm{w}_0 = \mathrm{v}_0$. From the discussion in Section 5.2.1, we see that this corresponds to using past observations over the same length of time, $t_S = t_Y$, for both S and Y in fixing the prior. Then $\mathrm{c}_0 = \rho \mathrm{v}_0$, and the solution to the system of equations (5.21) gives the entries of the matrix C_t as

$$\mathrm{v}_t = \frac{\mathrm{v}_0}{1 + \mathrm{v}_0 t}, \quad \mathrm{w}_t = \mathrm{v}_t, \quad \mathrm{c}_t = \rho \mathrm{v}_t.$$

With this simplification, the equation for the optimal filter simplifies to

$$\begin{pmatrix} \mathrm{d}\widehat{\lambda}_t \\ \mathrm{d}\widehat{\theta}_t \end{pmatrix} = \mathrm{v}_t \begin{pmatrix} \mathrm{d}\xi_t - \widehat{\lambda}_t \mathrm{d}t \\ \mathrm{d}\zeta_t - \widehat{\theta}_t \mathrm{d}t \end{pmatrix} = \mathrm{v}_t \begin{pmatrix} \mathrm{d}\widehat{B}_t \\ \mathrm{d}\widehat{W}_t \end{pmatrix},$$

which, along with the initial condition in (5.13), yields (5.16) and (5.17).

Finally, the expressions in (5.18) for ξ_t, ζ_t follow directly from the solutions of (5.1) and (5.2) for S and Y. □

Armed with Proposition 5.5 we may now treat the model as a full information model with random drift parameters $(\widehat{\lambda}_t, \widehat{\theta}_t)$, and this is done in the next section.

5.2.3 Optimal hedging with random drifts

On the stochastic basis $(\Omega, \widehat{\mathcal{F}}_T, \widehat{\mathbb{F}}, P)$, the wealth process associated with trading strategy $\pi := (\pi_t)_{0\leq t\leq T}$, an $\widehat{\mathbb{F}}$-adapted process satisfying $\int_0^T \pi_t^2 \mathrm{d}t < \infty$ a.s., is $X = (X_t)_{0\leq t\leq T}$, satisfying

$$\mathrm{d}X_t = \sigma\pi_t(\widehat{\lambda}_t \mathrm{d}t + \mathrm{d}\widehat{B}_t). \tag{5.22}$$

The class $\mathcal{M}$ of local martingale measures for this model consists of measures Q with density processes defined by

$$Z_t := \left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\widehat{\mathcal{F}}_t} = \mathcal{E}(-\widehat{\lambda}\cdot\widehat{B} - \psi\cdot\widehat{B}^{\perp})_t, \quad 0 \leq t \leq T, \tag{5.23}$$

for integrands ψ satisfying $\int_0^t \psi_s^2 \mathrm{d}s < \infty$ a.s., for all $t \in [0,T]$ (it is not hard to show that $\int_0^t \widehat{\lambda}_s^2 \mathrm{d}s < \infty, 0 \leq t \leq T$). For $\psi = 0$ we obtain the minimal martingale measure Q^M.

Under $Q \in \mathcal{M}$, $(\widehat{B}^Q, \widehat{B}^{\perp,Q})$ is two-dimensional Brownian motion, where

$$\mathrm{d}\widehat{B}_t^Q := \mathrm{d}\widehat{B}_t^Q + \widehat{\lambda}_t \mathrm{d}t, \qquad \mathrm{d}\widehat{B}_t^{\perp,Q} := \mathrm{d}\widehat{B}_t^{\perp} + \psi_t \mathrm{d}t,$$

and the asset prices and random drifts satisfy

$$\begin{array}{rcl} \mathrm{d}S_t &=& \sigma S_t \mathrm{d}\widehat{B}_t^Q, \\ \mathrm{d}Y_t &=& \beta Y_t[(\widehat{\theta}_t - \rho\widehat{\lambda}_t - \sqrt{1-\rho^2}\psi_t)\mathrm{d}t + \mathrm{d}\widehat{W}_t^Q], \\ \mathrm{d}\widehat{\lambda}_t &=& \mathrm{v}_t[-\widehat{\lambda}_t\mathrm{d}t + \mathrm{d}\widehat{B}_t^Q], \\ \mathrm{d}\widehat{\theta}_t &=& \mathrm{v}_t[-(\rho\widehat{\lambda}_t + \sqrt{1-\rho^2}\psi_t)\mathrm{d}t + \mathrm{d}\widehat{W}_t^Q], \end{array} \tag{5.24}$$

where $\widehat{W}^Q = \rho\widehat{B}^Q + \sqrt{1-\rho^2}\widehat{B}^{\perp,Q}$.

The relative entropy between $Q \in \mathcal{M}$ and P is defined by

$$\begin{array}{rcl} \mathcal{H}(Q,P) &:=& E\left[\dfrac{\mathrm{d}Q}{\mathrm{d}P}\log\dfrac{\mathrm{d}Q}{\mathrm{d}P}\right] \\ &=& E^Q\left[-\displaystyle\int_0^T \widehat{\lambda}_t \mathrm{d}\widehat{B}_t^Q - \int_0^T \psi_t \mathrm{d}\widehat{B}_t^{\perp,Q} + \frac{1}{2}\int_0^T \left(\widehat{\lambda}_t^2 + \psi_t^2\right)\mathrm{d}t\right]. \end{array}$$

Using the Q-dynamics of $\widehat{\lambda}_t$ it is straightforward to establish that $E^Q \int_0^t \widehat{\lambda}_s^2 \mathrm{d}s < \infty$ for all $t \in [0,T]$. If, in addition, we have the integrability condition

$$E^Q \int_0^t \psi_s^2 \mathrm{d}s < \infty, \quad 0 \le t \le T, \tag{5.25}$$

then

$$\mathcal{H}(Q,P) = E^Q \left[\frac{1}{2} \int_0^T \left(\widehat{\lambda}_t^2 + \psi_t^2 \right) \mathrm{d}t \right] < \infty. \tag{5.26}$$

In this case we write $Q \in \mathcal{M}_f$, where $\mathcal{M}_f$ denotes the set of martingale measures Q with finite relative entropy with respect to P, and we define $\mathcal{H}(Q,P) := \infty$ otherwise. From (5.26) we note that the minimal entropy measure Q^E is characterised by

$$\mathcal{H}(Q^E,P) = E^{Q^E} \left[\frac{1}{2} \int_0^T \widehat{\lambda}_t^2 \mathrm{d}t \right],$$

corresponding to $\psi \equiv 0$ in (5.26). This means that the minimal martingale measure and the minimal entropy measure in this model coincide: $Q^E = Q^M$.

For an initial time $t \in [0,T]$, we define the conditional entropy between $Q \in \mathcal{M}$ and P by

$$H_t(Q,P) := E \left[\frac{Z_T}{Z_t} \log \left(\frac{Z_T}{Z_t} \right) \middle| \widehat{\mathcal{F}}_t \right], \quad 0 \le t \le T, \tag{5.27}$$

satisfying $H_0(Q,P) \equiv \mathcal{H}(Q,P)$. Provided the integrability condition (5.25) is satisfied, then

$$H_t(Q,P) = E^Q \left[\frac{1}{2} \int_t^T \left(\widehat{\lambda}_u^2 + \psi_u^2 \right) \mathrm{d}u \middle| \widehat{\mathcal{F}}_t \right],$$

and we define $H_t(Q,P) := \infty$ otherwise. In particular, therefore, recalling that $\widehat{\lambda}_t \equiv \widehat{\lambda}(t,S_t)$ is a smooth and Lipschitz function of time and current stock price, and that the Q-dynamics of $\widehat{\lambda}_t$ do not depend on ψ_t for any $Q \in \mathcal{M}$, the minimal conditional entropy $(H_t(Q^E,P))_{0 \le t \le T}$ will be a deterministic function of time and stock price, given by $H_t(Q^E,P) \equiv H^E(t,S_t)$ for a $C^{1,2}([0,T] \times \mathbb{R}^+)$ function H^E defined by

$$H^E(t,s) := E^{Q^E} \left[\frac{1}{2} \int_t^T \widehat{\lambda}^2(u,S_u) \mathrm{d}u \middle| S_t = s \right]. \tag{5.28}$$

5.2.4 The primal problem

We use an exponential utility function, $U(x) = -\exp(-\alpha x)$, $x \in \mathbb{R}$, $\alpha > 0$. The primal value function $u^{(n)}$ is defined as the maximum expected utility of wealth at T from trading S and receiving n units of the claim on Y, when starting at time $t \in [0,T]$:

$$u^{(n)}(t,x,s,y) := \sup_{\pi \in \mathcal{A}} E\big[U(X_T + nh(Y_T)) \mid X_t = x, S_t = s, Y_t = y\big], \tag{5.29}$$

where $\mathcal{A}$ denotes the set of admissible trading strategies. The dynamics of the state variables X, S, Y are given by (5.22) and (5.14, 5.15). For starting time zero we write $u^{(n)}(x) \equiv u^{(n)}(0, x, \cdot, \cdot)$.

The set of admissible strategies is defined as follows. Denote by $\Delta := \pi/S$ be the adapted S-integrable process for the number of shares held. The space of permitted strategies is

$$\mathcal{A} = \{\Delta : (\Delta \cdot S) \text{ is a } (Q, \widehat{\mathbb{F}})\text{-martingale for all } Q \in \mathcal{M}_f\},$$

where $(\Delta \cdot S)_t = \int_0^t \Delta_u \mathrm{d}S_u$ is the gain from trading over $[0,t], t \in [0,T]$.

Denote the optimal trading strategy by $\pi^* \equiv \pi^{*,n}$, and the optimal wealth process by $X^* \equiv X^{*,n}$. The utility-based price $p^{(n)}$ and optimal hedge for a position in n claims are defined along the lines of Definitions 5.1 and 5.2. The indifference price per claim at $t \in [0,T]$, given $X_t = x, S_t = s, Y_t = y$, is $p^{(n)}$ given by

$$u^{(n)}(t, x - np^{(n)}(t,x,s,y), s, y) = u^{(0)}(t,x,s).$$

The optimal hedging strategy is to hold $(\Delta_t^H)_{0 \le t \le T}$ shares of stock at time t, where $\Delta_t^H S_t =: \pi_t^H S_t$, and $\pi^H := \left(\pi_t^H\right)_{0 \le t \le T}$, is defined by

$$\pi_t^H := \pi_t^{*,n} - \pi_t^{*,0}, \quad 0 \le t \le T. \tag{5.30}$$

It is well known that with exponential utility the indifference price is independent of the initial cash wealth x, so we shall write $p^{(n)}(t,x,s,y) \equiv p^{(n)}(t,s,y)$ from now on.

For small positions in the claim (or, equivalently, for small risk aversion), we shall later approximate the indifference price by the marginal utility-based price introduced by Davis [3]. This is the indifference price for infinitesimal diversions of funds into the purchase or sale of claims, and is equivalent (as is well-known, see for example Monoyios [24]) to the limit of the indifference price as $n \to 0$.

Definition 5.6 (Marginal price). The marginal utility-based price of the claim at $t \in [0,T]$ is $\widehat{p}(t,s,y)$ defined by

$$\widehat{p}(t,s,y) := \lim_{n \to 0} p^{(n)}(t,s,y).$$

It is well known that with exponential utility the marginal price is also equivalent to the limit of the indifference price as risk aversion goes to zero. Under appropriate conditions (satisfied in this model) it is given by the expectation of the payoff under the optimal measure of the dual problem without the claim. For exponential utility this measure is the minimal entropy measure Q^E and, as we have already seen, in our model $Q^E = Q^M$, giving the representation $\widehat{p}(t,s,y) = E^{Q^M}\left[h(Y_T) \mid S_t = s, Y_t = y\right]$, as we shall see in the next section.

5.2.5 Dual problem and optimal hedge

We attack the primal utility maximisation problem (5.29) using classical duality results. For a problem with the random terminal endowment of a European claim, and with exponential utility, as in this paper, Delbaen et al. [5] establish the required duality relations between the primal and dual problems in a semimartingale setting. We shall use these results below to establish a simple algebraic relation (Lemma 5.7) between the primal value function and the indifference price, which we shall then exploit to derive the representation for the optimal hedging strategy.

The dual problem with starting time zero has value function defined by

$$\tilde{u}^{(n)}(\eta) := \inf_{Q\in\mathcal{M}} E\big[\widetilde{U}(\eta Z_T) + \eta Z_T n h(Y_T)\big],$$

where Z is the density process in (5.23) and $\widetilde{U}$ is the convex conjugate of the utility function. For exponential utility $\widetilde{U}$ is given by

$$\widetilde{U}(\eta) = \frac{\eta}{\alpha}\left[\log\left(\frac{\eta}{\alpha}\right) - 1\right].$$

Hence the dual value function has the well-known entropic representation

$$\tilde{u}^{(n)}(\eta) = \widetilde{U}(\eta) + \frac{\eta}{\alpha}\inf_{Q\in\mathcal{M}}\left[\mathcal{H}(Q,P) + \alpha n E^Q h(Y_T)\right].$$

Denoting the dual minimiser that attains the above infimum by $Q^{*,n}$, we observe that $Q^{*,n}\in\mathcal{M}_f$.

For a starting time $t\in[0,T]$ the dual value function is defined by

$$\tilde{u}^{(n)}(t,\eta,s,y) := \inf_{Q\in\mathcal{M}} E\left[\widetilde{U}\left(\eta\frac{Z_T}{Z_t}\right) + \eta\frac{Z_T}{Z_t} n h(Y_T)\middle| S_t = s, Y_t = y\right], \tag{5.31}$$

and we write $\tilde{u}^{(n)}(\eta) \equiv \tilde{u}^{(n)}(0,\eta,\cdot,\cdot)$.

Lemma 5.7. *The primal value function and indifference price are related by*

$$u^{(n)}(t,x,s,y) = u^{(0)}(t,x,s)\exp\left(-\alpha n p^{(n)}(t,s,y)\right), \tag{5.32}$$

where the value function without the claim is given by

$$u^{(0)}(t,x,s) = -\exp\left(-\alpha x - H^E(t,s)\right), \tag{5.33}$$

and $H^E(t,s)$ is the conditional minimal entropy function defined in (5.28).

Proof. For brevity, we give the proof for $t=0$. The proof for a general starting time follows similar lines, and we make some comments on how to adapt the following argument for that case at the end of the proof.

The fundamental duality linking the primal and dual problems in Delbaen et al. [5] implies that the value functions $u^{(n)}(x)$ and $\tilde{u}^{(n)}(\eta)$ are conjugate:

$$\tilde{u}^{(n)}(\eta) = \sup_{x\in\mathbb{R}}[u^{(n)}(x) - x\eta], \qquad u^{(n)}(x) = \inf_{\eta>0}[\tilde{u}^{(n)}(\eta) + x\eta].$$

The value of η attaining the above infimum is η^*, given by $\tilde{u}^{(n)}_\eta(\eta^*) = -x$, so that

$$u^{(n)}(x) = \tilde{u}^{(n)}(\eta^*) + x\eta^*,$$

which translates to

$$u^{(n)}(x) = -\exp\left(-\alpha x - \inf_{Q\in\mathcal{M}}\left[\mathcal{H}(Q,P) + \alpha n E^{Q} h(Y_T)\right]\right). \tag{5.34}$$

So, in particular,

$$u^{(0)}(x) = -\exp\left[-\alpha x - \mathcal{H}(Q^E, P)\right], \tag{5.35}$$

where Q^E is the minimal entropy measure: $Q^E = Q^{*,0}$.

Combining the dual representations (5.34) and (5.35) for the primal problems with and without the claim, with the definition of the indifference price, gives the dual representation for the utility-based price in the form

$$p^{(n)} = \frac{1}{\alpha n}\left[\inf_{Q\in\mathcal{M}}\left[\mathcal{H}(Q,P) + \alpha n E^{Q} h(Y_T)\right] - \mathcal{H}(Q^E,P)\right], \tag{5.36}$$

which is the representation found in Delbaen et al. [5], modified slightly as we have a random endowment of n claims ([5] considered the case $n = -1$).

In particular, for $n \to 0$ or $\alpha \to 0$, we obtain the marginal price of Davis [3]:

$$\widehat{p} := \lim_{n\to 0} p^{(n)} = E^{Q^E} h(Y_T) = E^{Q^M} h(Y_T), \tag{5.37}$$

the last equality following from the equality of Q^M and Q^E, as implied by (5.26).

From (5.34)–(5.36), the relation between the primal value functions and indifference price then follows immediately, as

$$\begin{aligned} u^{(n)}(x) &= -\exp\left(-\alpha x - \mathcal{H}(Q^E,P) - \alpha n p^{(n)}\right) \\ &= u^{(0)}(x)\exp\left(-\alpha n p^{(n)}\right). \end{aligned}$$

Similarly, a corresponding relation for a starting time $t \in [0,T]$ may also be derived. This is achieved using the definition (5.31) of the dual value function for an initial time $t \in [0,T]$, the conjugacy of $u^{(n)}(t,x,s,y)$ and $\tilde{u}^{(n)}(t,\eta,s,y)$ and the definitions (5.27) and (5.28) of the conditional entropy and conditional minimal entropy. □

Using Lemma 5.7 we obtain the following representation for the optimal hedging strategy associated with the indifference price. In what follows we assume that the indifference price is a suitably smooth function of (t,s,y), so that (given Lemma 5.7) we may assume the primal value function is smooth enough to be a classical solution of the associated Hamilton–Jacobi–Bellman (HJB) equation. This smoothness property is confirmed in [23].

Theorem 5.8. *The optimal hedge for a position in n claims is to hold Δ_t^H units of S at $t \in [0,T]$, where*

$$\Delta_t^H = -n\left(p_s^{(n)}(t,S_t,Y_t) + \rho\frac{\beta}{\sigma}\frac{Y_t}{S_t}p_y^{(n)}(t,S_t,Y_t)\right).$$

Remark 5.9. We note the extra term in the hedging formula compared with the corresponding full information result (5.10). The drift parameter uncertainty results in additional risk, manifested as dependence of the indifference price on the stock price, and hence the derivative with respect to the stock price appears in the theorem.

Proof. The HJB equation associated with the primal the value function is

$$u_t^{(n)} + \max_{\pi} \mathcal{A}_{X,S,Y} u^{(n)} = 0,$$

where $\mathcal{A}_{X,S,Y}$ is the generator of (X, S, Y) under P. Performing the maximisation over π yields the optimal Markov control as $\pi_t^{*,n} = \pi^{*,n}(t, X_t^{*,n}, S_t, Y_t)$, where

$$\pi^{*,n}(t,x,s,y) = -\left(\frac{\widehat{\lambda} u_x^{(n)} + \sigma s u_{xs}^{(n)} + \rho\beta y u_{xy}^{(n)}}{\sigma u_{xx}^{(n)}}\right),$$

and where the arguments of the functions on the right-hand side are omitted for brevity. For the case $n = 0$ there is no dependence on y in the value function $u^{(0)}$, and we have $\pi_t^{*,0} = \pi^{*,0}(t, X_t^{*,0}, S_t)$, where

$$\pi^{*,0}(t,x,s) = -\left(\frac{\widehat{\lambda} u_x^{(0)} + \sigma s u_{xs}^{(0)}}{\sigma u_{xx}^{(0)}}\right).$$

Applying the definition (5.30) of the optimal hedging strategy along with the representations (5.32) and (5.33) from Lemma 5.7 for the value functions, gives the result. □

5.2.6 Stochastic control representation of the indifference price

Using the expression (5.26) for the relative entropy between measures in $Q \in \mathcal{M}_f$ and P in the dual representation (5.36) of $p^{(n)}$, we obtain the indifference price of the claim at time zero as the value function of a control problem:

$$p^{(n)} = \inf_{\psi} E^Q\left[\frac{1}{2\alpha n}\int_0^T \psi_t^2 \mathrm{d}t + h(Y_T)\right],$$

to be minimised over control processes $(\psi_t)_{0\leq t\leq T}$, such that $Q \in \mathcal{M}_f$. Of course, we need only consider measures with finite relative entropy since a martingale measure with $\mathcal{H}(Q,P) = \infty$ will not attain the infimum in (5.36). The dynamics for S, Y are given in the system of equations (5.24). Equivalently, since $\widehat{\lambda}, \widehat{\theta}$ may be expressed as functions of time and current asset price by (5.17), we may write the state dynamics of the control problem for the indifference price as

$$\begin{aligned} \mathrm{d}S_t &= \sigma S_t \mathrm{d}\widehat{B}_t^Q, \\ \mathrm{d}Y_t &= \beta Y_t[(\widehat{\theta}(t,Y_t) - \rho\widehat{\lambda}(t,S_t) - \sqrt{1-\rho^2}\psi_t)\mathrm{d}t + \mathrm{d}\widehat{W}_t^Q]. \end{aligned}$$

Adopting a dynamic programming approach, we consider a starting time $t \in [0, T]$. Then we have

$$p^{(n)}(t,s,y) = \inf_{\psi} E^Q\left[\frac{1}{2\alpha n}\int_t^T \psi_u^2 \mathrm{d}u + h(Y_T)\middle| S_t = s, Y_t = y\right].$$

The HJB dynamic programming PDE associated with $p^{(n)}(t,s,y)$ is

$$p_t^{(n)} + \mathcal{A}_{S,Y}^{Q^M} p^{(n)} + \inf_{\psi}\left[\frac{1}{2\alpha n}\psi^2 - \beta\sqrt{1-\rho^2}\psi y p_y^{(n)}\right] = 0, \quad p(T,s,y) = h(y),$$

where $\mathcal{A}_{S,Y}^{Q^M}$ is the generator of (S,Y) under the minimal measure:

$$\mathcal{A}_{S,Y}^{Q^M} f(t,s,y) = \beta(\widehat{\theta}(t,y) - \rho\widehat{\lambda}(t,s))yf_y + \frac{1}{2}s^s f_{ss} + \frac{1}{2}\beta^2 y^2 f_{yy} + \rho\sigma\beta sy f_{sy}.$$

Performing the minimisation in the HJB equation, the optimal Markov control is $\psi_t^{*,n} \equiv \psi^{*,n}(t,S_t,Y_t)$, where

$$\psi^{*,n}(t,s,y) = \alpha n\sqrt{1-\rho^2}\beta y p_y^{(n)}(t,s,y),$$

and note that $\psi^{*,0} = 0$. Substituting back into the HJB equation, we find that $p^{(n)}$ solves the semi-linear PDE

$$p_t^{(n)} + \mathcal{A}_{S,Y}^{Q^M} p^{(n)} - \frac{1}{2}\alpha n(1-\rho^2)\beta^2 y^2 \left(p_y^{(n)}\right)^2 = 0, \quad p^{(n)}(T,s,y) = h(y).$$

We note that for $n = 0$ this becomes a linear PDE for the marginal price $\widehat{p}$, so that by the Feynman–Kac theorem we have

$$\widehat{p}(t,s,y) = E_{t,s,y}^{Q^M} h(Y_T), \tag{5.38}$$

consistent with the general result (5.37). We shall see that in this case the marginal price is given by a BS-type formula.

5.2.7 Analytic approximation for the indifference price

To obtain analytic results we approximate the indifference price by the marginal price in (5.38). The marginal price (and hence the associated trading strategy) can be computed in analytic form since, under Q^M, $\log Y_T$ is Gaussian. We have the following result.

Proposition 5.10. *Under Q^M, conditional on $S_t = s, Y_t = y$, $\log Y_T \sim \mathrm{N}(m,\Sigma^2)$, where $m \equiv m(t,s,y)$ and $\Sigma^2 \equiv \Sigma^2(t)$ are given by*

$$\begin{aligned} m(t,s,y) &= \log y + \beta\left(\widehat{\theta}(t,y) - \rho\widehat{\lambda}(t,s) - \frac{1}{2}\beta\right)(T-t) \\ \Sigma^2(t) &= \left[1 + (1-\rho^2)\mathrm{v}_t(T-t)\right]\beta^2(T-t)\,. \end{aligned}$$

Proof. This is established by computing the SDEs for Y and for $\widehat{\theta}_t - \rho\widehat{\lambda}_t$ under Q^M. Indeed, applying the Itô formula to $\log Y_t$ under Q^M, we obtain, for $t < T$,

$$\log Y_T = \log Y_t + \beta\int_t^T \left(\widehat{\theta}_u - \rho\widehat{\lambda}_u\right)\mathrm{d}u - \frac{1}{2}\beta^2(T-t) + \beta\int_t^T \mathrm{d}\widehat{W}_u^{Q^M}, \tag{5.39}$$

where $\widehat{W}^{Q^M}$ is a Brownian motion under Q^M. The dynamics of $\widehat{\theta}_t - \rho\widehat{\lambda}_t$ under Q^M are

$$\mathrm{d}(\widehat{\theta}_t - \rho\widehat{\lambda}_t) = \sqrt{1-\rho^2}\mathrm{v}_t \mathrm{d}\widehat{B}_t^{\perp,Q^M},$$

where $\widehat{B}^{\perp,Q^M}$ is a Q^M-Brownian motion perpendicular to that driving the stock, related to $\widehat{W}^{Q^M}$ by $\widehat{W}^{Q^M} = \rho\widehat{B}^{Q^M} + \sqrt{1-\rho^2}\widehat{B}^{\perp,Q^M}$, and where $\widehat{B}^{Q^M}$ is the Brownian motion driving S. Hence, for $u > t$, after changing the order of integration in a double integral, we obtain

$$\int_t^T \left(\widehat{\theta}_u - \rho\widehat{\lambda}_u\right)\mathrm{d}u = \left(\widehat{\theta}_t - \rho\widehat{\lambda}_t\right)(T-t) + \sqrt{1-\rho^2}\int_t^T \mathrm{v}_u(T-u)\mathrm{d}\widehat{B}_u^{\perp,Q^M}.$$

This can be inserted into (5.39) to yield the desired result. □

We are thus able to obtain BS-style formulae for the price and hedge. For a put option of strike K we easily obtain the following explicit formulae for the marginal price and the associated optimal hedging strategy, where Φ denotes the standard cumulative normal distribution function.

Corollary 5.11. *With m and Σ as in Proposition 5.10, define $b \equiv b(t,s,y)$ by*

$$m = \log y + b - \frac{1}{2}\Sigma^2.$$

Then the marginal price at time $t \in [0,T]$ of a put option with payoff $(K - Y_T)^+$ is $\widehat{p}(t, S_t, Y_t)$, where

$$\begin{aligned} \widehat{p}(t,s,y) &= K\Phi(-d_1 + \Sigma) - ye^b\Phi(-d_1), \\ d_1 &= \frac{1}{\Sigma}\left[\log\left(\frac{y}{K}\right) + b + \frac{1}{2}\Sigma^2\right]. \end{aligned}$$

The optimal hedging strategy given by Theorem 5.8 with $\widehat{p}$ as an approximation to the indifference price is $\widehat{\Delta}_t \equiv \widehat{\Delta}(t, S_t, Y_t)$, where

$$\widehat{\Delta}(t,s,y) = n\rho\frac{\beta}{\sigma}\frac{y}{s}e^b\Phi(-d_1).$$

In [23] these results are used to conduct a simulation study of the effectiveness of the optimal hedge under partial information (that is, with Bayesian learning about the drift parameters of the assets), compared with the BS-style hedge and the optimal hedge without learning. The results show that optimal hedging combined with a filtering algorithm to deal with drift parameter uncertainty can indeed give improved hedging performance over methods which take S as a perfect proxy for Y, and over methods which do not incorporate learning via filtering.

Bibliography

[1] T. Björk, M. H. A. Davis and C. Landén, *Optimal investment under partial information*, preprint (2008)

[2] A. Danilova, M. Monoyios and A. Ng, *Optimal investment with inside information and parameter uncertainty*, preprint (2009).

[3] M. H. A. Davis, *Option pricing in incomplete markets*, Mathematics of Derivative Securities (Cambridge) (M. A. H. Dempster and S. R. Pliska, eds.), Cambridge University Press, 1997, pp. 216–226.

[4] M. H. A. Davis, *Optimal hedging with basis risk* From Stochastic Calculus to Mathematical Finance : The Shiryaev Festschrift (Berlin) (Y. Kabanov, R. Lipster and J. Stoyanov, eds.), Springer, 2006, pp. 169–188.

[5] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer and C. Stricker, *Exponential hedging and entropic penalties*, Mathematical Finance 12 (2002), pp. 99–123.

[6] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, 1975.

[7] M. Fujisaki, G. Kallianpur and H. Kunita, *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math. 9 (1972), pp. 19–40.

[8] V. Henderson, *Valuation of claims on nontraded assets using utility maximization*, Mathematical Finance 12 (2002), pp. 351–373.

[9] J. Hugonnier and D. Kramkov, *Optimal investments with random endowments in incomplete markets*, Annals of Applied Probability 14 (2004), pp. 845–864.

[10] T. Kailath, *An innovations approach to least-squares estimation, Part I: Linear Filtering in Additive Noise*, IEEE Trans. Automatic Control 13 (1968), pp. 646–655.

[11] R. E. Kalman and R. S. Bucy, *New results in linear filtering and prediction theory*, Trans. ASME Ser. D. J. Basic Engineering 83 (1961), pp. 95–108.

[12] G. Kallianpur, *Stochastic Filtering Theory*, Springer, 1980.

[13] I. Karatzas, I *Lectures on the Mathematics of Finance*, CRM Monographs 8 (1996), American Mathematical Society.

[14] I. Karatzas, J. P. Lehoczky, S. E. Shreve and G-L. Xu, 1991 *Martingale and duality methods for utility maximization in an incomplete market*, SIAM Journal on Control and Optimization 29 (1991), pp. 702–730.

[15] D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, Annals of Applied Probability 9 (1999), pp. 904–950.

[16] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes I: General Theory*, Springer, 2001.

[17] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes II: Applications*, Springer, 2001.

[18] R. Mansuy and M. Yor, *Random times and enlargements of filtrations in a Brownian setting*, Lecture Notes in Mathematics 1873 (2006), Springer.

[19] R. C. Merton, *Lifetime portfolio selection under uncertainty: the continuous-time case*, Rev. Econom. Statist. 51 (1969), pp. 247–257.

[20] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J. Econom. Theory 3 (1971), pp. 373–413. *Erratum:* ibid. 6 (1973) pp. 213–214.

[21] M. Monoyios, *Performance of utility-based strategies for hedging basis risk*, Quantitative Finance 4 (2004), pp. 245–255.

[22] M. Monoyios, *Optimal hedging and parameter uncertainty*, IMA Journal of Management Mathematics 18 (2007), pp. 331–351.

[23] M. Monoyios, *Marginal utility-based hedging of claims on non-traded assets with partial information*, preprint (2008).

[24] M. Monoyios, *Utility indifference pricing with market incompleteness*, Nonlinear Models in Mathematical Finance: New Research Trends in Option Pricing (New York) (M. Ehrhardt, ed.), Nova Science Publishers, 2008.

[25] M. Monoyios, *Asymptotic expansions for optimal hedging of basis risk with partial information*, preprint (2009).

[26] M. Musiela and T. Zariphopoulou, *An example of indifference prices under exponential preferences*, Finance & Stochastics 8 (2004), pp. 229–239.

[27] M. P. Owen, *Utility based optimal hedging in incomplete markets*, Annals of Applied Probability 12 (2002), pp. 691–709.

[28] H. Pham, *Portfolio optimization under partial observation: theoretical and numerical aspects*, Handbook of Nonlinear Filtering (Oxford) (D. Crisan and B. Rozovsky, eds.), Oxford University Press, to appear.

[29] I. Pikovsky and I. Karatzas, *Anticipative portfolio optimization*, Adv. in App. Prob. 28 (1996), pp. 1095–1122.

[30] E. Platen and W. J. Runggaldier, *A benchmark approach to portfolio optimization under partial information*, Asia Pacific Financial Markets 14 (2007), pp. 25–43.

[31] L. C. G. Rogers, *The relaxed investor and parameter uncertainty*, Finance & Stochastics 5 (2001), pp. 131–154.

[32] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales, Vol. 2: Itô Calculus*, 2nd. ed., Cambridge University Press, Cambridge, UK, 2000.

[33] M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 11 (1969), pp. 230–243.

Author information

Michael Monoyios, Mathematical Institute, University of Oxford, UK.
Email: monoyios@maths.ox.ac.uk